AFOSR-TR-97-0474

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services. Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget. Paperwork Reduction Project (0704-0188) Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1997	Final, October 1993- September 1996

4. TITLE AND SUBTITLE

Robust Multiresolution Integrated Target Sensing and Recognition

5. FUNDING NUMBERS

AFOSR-F49620-93-1-0558

6. AUTHOR(S)

Dr. A. H. Tewfik

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(S)

Dept. of Electrical Engineering
University of Minnesota
Room 4-174 EE/CSci Bldg
Minneapolis, MN 55455

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)

Dr. Jon Sjogren,
AFOSR/NM
Bldg. 410
Bolling Air Force Base, Washington, D.C.

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION UNLIMITED

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The main goal of this research effort was to develop an integrated target sensing and recognition strategy. Secondary goals of this work were to construct novel image representation and analysis algorithms to facilitate content based image retrieval. We derived a new adaptive waveform selection algorithm for radar range-Doppler target recognition. The algorithm selects the waveforms that provide maximum discrimination information at any given time by maximizing the Kullback-Leibler information number corresponding to the most likely hypothesis. As a result, it minimizes decision time for a given desired classification performance level and maximizes classification performance for a fixed data acquisition time. We have developed a novel theory of multiresolution representation of binary data that supports fast binary image matching algorithms. Only Boolean operation are needed to compute these representations. Finally, we have constructed a novel image coding technique that supports pictorial queries. The procedure minimizes a weighted sum of the expected compressed image file size in bits and the expected number of bits that need to be read to answer a pictorial query (query response time).

14. SUBJECT TERMS

target recognition, target sensing, wavelets, radar, SAR, ISAR, beamforming, adaptive algorithms, array, image databases, pictorial queries, image coding, multiresolution analysis

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (REV. 2-89)
Prescribed by ANSI Std. Z39-18
296-132

DTIC QUALITY INSPECTED 3

Executive Summary

The main goal of this research effort is to develop an integrated target sensing and recognition strategy. Secondary goals of this work are to construct novel image representation and analysis algorithms to facilitate content based image retrieval.

We apply a sequential experiment design procedure to the problem of signal selection for radar target classification. In this approach, radar waveforms are designed to discriminate between targets possessing a doubly-spread *reflectivity function* that are observed in clutter. The waveforms minimize decision time by maximizing the discrimination information in the echo signal. Each selected waveform maximizes the *Kullback-Leibler information number* that measures the dissimilarity between the observed target and alternative targets. We discuss in details two scenarios. In the first scenario, the target environment is assumed fixed during illumination. In this case the optimal waveform selection strategy leads to a fixed library of waveforms. During actual classification, the sequence in which the waveforms are selected from the library is determined from the noise to clutter power in the range-Doppler support of the targets. In the second scenario, the target environment changes between pulse transmissions. In this case, the maximum discrimination information is obtained by a repeated transmission of a single waveform designed from the reflectivity function of the targets. We show that our choice of signals can produce significant gains in detection performance.

We develop a new image representation which combines support for coding and content-based retrieval. The algorithm minimizes a weighted sum of the expected compressed file size and query response time. Our approach leads to a progressive refinement retrieval by successively reducing the number of searched files as more bits are read. Furthermore, no distance computation is required during a query. Only simple bit pattern comparisons are required. The approach supports compressed data modification and low-bit-rate high quality browsing.

Finally, we construct a theory of binary wavelet decompositions of finite binary images. The new binary wavelet transform uses simple modulo-2 operations. It shares many of the important characteristics of the real wavelet transform. In particular, it yields an output similar to the thresholded output of a real wavelet transform operating on the underlying binary image. We introduce a new binary field transform to use as an alternative to the discrete Fourier transform over $GF(2)$. The corresponding concept of sequence spectra over $GF(2)$ is defined. Using this transform, a theory of binary wavelets is developed in terms of 2-band perfect reconstruction filter banks in $GF(2)$. By generalizing the corresponding real field constraints of bandwidth, vanishing moments and spectral content in the filters, we cconstruct a perfect reconstruction wavelet decomposition. We also demonstrate the potential use of the binary wavelet decomposition in lossless image coding.

Contents

Chapter 1

Introduction

Today an automatic target recognition (ATR) system designer has a wide array of sensors to choose from. These include active sensors (e.g. millimeter and laser radars) and passive sensors (e.g. passive sonar, forward looking infrared (FLIR) systems). The price of these sensors can vastly increase the cost of the target recognition system. Some of these sensors could reveal the presence of the ATR system to an enemy. Furthermore, any sensor in the system can collect a large amount of information in short times. The size of this information can easily overwhelm the capabilities of today's most powerful computers. Yet only a fraction of the collected information may be relevant to the recognition task. Furthermore, an ATR system should make its classification decisions in real time. This of course is related to the computational complexity of the algorithm used as well as to the ability of the system to examine and seek only relevant information. Finally, an ATR system must perform well under a wide variety of scenarios. In particular, its performance should not degrade dramatically when it is used in an environment or domain for which it was not initially trained. Thus an ATR system can be characterized by four parameters: the amount of resources that it uses (sensors and computers), its computational complexity as measured by the time it takes to make a decision, its robustness to changes in the environment and its performance as described by its probability of correctly classifying a target (probability of detection) versus the probability of mistaking another type of targets for the desired type (probability of false alarm). *A good ATR system is one that can produce accurate classification decisions in a minimum time while expanding a minimum amount of resources in a wide variety of circumstances. Its performance should degrade gracefully (rather than catastrophically) in adverse environments.* The goal of this research program was to construct a methodology for designing such good ATR system. The methodology is general, i.e., it is applicable to a wide class of ATR and other pattern recognition tasks. However, we demonstrated the methodology using a particular choice of sensors. Specifically, we considered radar sensors and showed how recognition can be improved (in terms of performance and the time it takes to reach a classification decision) by dynamically selecting the radar waveforms.

Our approach is to use an *intelligent multiresolution integrated sensing and classification strategy* . An intelligent and integrated strategy implies that there is feedback from the classification module to the sensing module. The feedback specifies which waveforms the radar ought to transmit. It means that the classification module uses a robust computationally efficient classification approach that adapts to the current environment and the target at hand. This adaptation is also used to control the data acquisition chore. Specifically, the algorithm uses its current classification of the target and information about the state of the environment to decide what data to collect next. It produces a sequence of progressively finer classification decisions. Relevant data is collected only when it is needed. Efficient data representations are used. Note that our approach to ATR may be best described as a robust adaptive approach that minimizes computational complexity while preserving a given classification performance. It goes *well beyond most adaptive* ATR algorithms that have been proposed recently. These algorithms change only certain parameters of the algorithm (e.g. the threshold at which an edge is detected). By contrast, our procedure uses an estimate of the state of the environment and its current knowledge about the target to select the next set of data that should be acquired. This is akin to changing matched filters as we collect more statistical information about a desired signal and an additive colored noise component. Furthermore, our procedure specifically tries to

minimize complexity while preserving a given performance measure.

Observe also that effective ATR necessarily builds on prior knowledge of the system. In particular, a good ATR system must be able to correlate the information that it receives with its own prior domain knowledge. In this project, we have focused on efficient querying of prior visual information (stored as image data) by pictorial content. We developed image coding techniques that minimize the bit size of the coded image and the retrieval time in response to a pictorial query. We have also developed a new binary multiresolution representation of binary images that is amenable to fast matching. The importance of this representation and the corresponding image matching algorithms is that they require only fast Boolean operations.

In the remainder of this report we describe in details the results that we have obtained. We have grouped these results under three headings:

1. Waveform selection for radar target classification. This part of our research is described in Chapter 2. The emphasis here is on integrated classification and data acquisition. The approach detailed in this chapter easily generalizes to multi-sensor systems with various degrees of freedom.

2. Image coding for query by pictorial content. Queries in image based ATR systems are pictorial in nature. Furthermore, ATR systems may have at their disposition large collections of images. In Chapter 3, we describe a novel approach for coding images that supports fast retrieval by pictorial content.

3. Binary wavelet decomposition of binary images. Images are often represented in binary form. We describe in Chapter 4 a new binary multiresolution representation of binary images that is amenable to fast matching.

Chapter 2

Waveform Selection for Radar Target Classification

2.1 Introduction

In this chapter we consider optimal signal design for the problem of radar target classification. The problem of radar signal design for classification is to find signals that discriminate between a collection of targets of interest by observing the backscatter from an illuminated unknown target. The objective is to find probing signals that enhance certain unique aspects of a particular target that distinguishes it from other targets. In order to assess the performance of a particular signal set, it is necessary to define a measure of goodness. For a classification problem, this measure is usually the probability of misclassification. A signal set is *better* than another in a given scenario, if it results in a smaller probability of misclassification. Very little has been done in the design of radar signals for target classification with the objective of minimizing the probability of misclassification. This is mainly due to the fact that it is generally difficult to obtain analytic expressions for the probability of misclassification that would indicate a favorable choice of one signal selection strategy over another. Employing techniques from sequential experiment design, we present a signal selection procedure for sequentially classifying targets that possess a doubly spread scattering function. A measure related to the probability of misclassification is used to assess the performance of the signal selection scheme. Signals are then selected to optimize this measure of dissimilarity.

The echo signal that backscatters from the target contains information about the target environment, shape and motion. Since we seek to identify special targets, information related to the target environment such as clutter and noise is treated as interference. Target information appears in the echo in the form of amplitude and scale variation of the transmitted signal [1]. These variations are caused by the reflective properties of the target and the relative motion between the radar platform and each point on the target surface. The reflection from these surfaces can be described by the two-dimensional *reflectivity* or *scattering function* of the target [2, 3, 4]. When a particular signal illuminates a target, the return will be a projection of the reflectivity function onto the one-dimensional subspace spanned by the dilates of the transmitted signal. Therefore, only limited target information can be obtained from the echo signal. To develop an effective signal selection strategy for classification, it is essential to obtain signals that can extract distinguishing properties from the observed target. Unlike current techniques that separate between the problem of waveform design and classification, we seek to design waveforms suited to a particular classification task and with the intention of minimizing the probability of misclassification.

There are two distinct approaches to target classification. The first is a nonparametric approach. In this approach, the backscatter from the targets of interests are experimentally or numerically determined. The pulse used to illuminate the targets is in general a short pulse and is unrelated to the ensemble of targets. Relevant features are then extracted from the backscatter signals of each target return and stored. Standard pattern recognition can then be applied by comparing the features of the observed target return to the stored target features and selecting the best match according to some distance criterion [5]. The nonparametric approach has been widely applied in practice for both standard and high resolution radar. In polarimetric

techniques for example, certain parameters (features) are extracted from the polarization scattering matrix and are used to design a standard pattern recognizer [6],[7],[8]. Other nonparametric methods include the singularity expansion methods. These methods are based on the fact that at low frequencies a target can be described by a transfer function with several resonant frequencies [9]. These resonant frequencies determine the size and shape of the target [10]. Various discrimination waveforms, synthesized to identify a specific target response from these natural frequencies have emerged. These are linear filters which, when convolved with the target responses to which they are matched, annihilate preselected natural-frequency content of those responses. The natural frequencies of the relevant target can be measured in the laboratory using scale-model targets. Some of the suggested waveforms include the K-pulse, the E-pulse and the S-pulse. For an excellent discussion of these discriminant waveforms, see [11] and references therein. In the parametric approaches, the underlying target and clutter statistics are assumed to be known. In this case, the decision process can be posed as a test of statistical hypotheses. The target models are obtained from the scattering behavior of isolated centers such as flat plates and corner reflectors that dominate a wideband radar return [12]. The required threshold is set to minimize the probability of error.

In contrast to the above approaches, we design a radar signal set to maximize classification performance. The resulting signal selection problem can essentially be treated in the context of experimental design [13]. Experiment design is an organized method for extracting as much information as possible from a limited number of observations. The number of observations need not be fixed and generally it is desirable to minimize the number of experiments that need to be performed to reach a decision. This leads to a sequential approach in experiment design.

Here, we propose to use a sequential classification procedure that minimizes the average number of necessary signal transmissions. Instead of minimizing the probability of error, we use the *Kullback-Leibler information numbers* [14] as distance measures between the probability density functions of the observations. These measures are used in information theory [65] to obtain bounds on the probability of error through communication channels and in statistics to obtain asymptotic bounds on Bayes' risk [16] in a binary hypothesis testing problem. Our approach to signal design is derived from the sequential experiment design due to Chernoff [16]. In this approach, an experiment is selected based on maximizing one of the Kullback-Leibler information numbers at each decision stage. The resulting procedure is sequential where the classifier attempts to reach a decision after each echo is received. It is optimal in the sense that it minimizes the average number of observations used by the sequential test provided that the observations are independent and identically distributed. It is also asymptotically optimum in the sense that if the number of observations is sufficiently large, the probability of error is minimized provided again that the observations are independent and identically distributed. When the target environment remains unchanged during illumination, the proposed method finds a sequence of pulses that in general have different shapes. The sequence in which the signals are transmitted is determined *offline* and is calculated based on the nature and type of the interference that dominates the return signal. If the interference comes mainly from the target environment, such as clutter, then our procedure selects orthogonal signals that provide discrimination information along new directions of the target reflectivity function. This is because clutter interference is assumed to remain fixed during illumination and is statistically the same in all signal directions. On the other hand, when interference comes mainly from observation noise, the procedure selects repeated signal transmissions to improve the quality of our measurements by averaging. When the target environment changes between pulse transmissions, the procedure selects only one optimum signal. It is in a sense a "universal" signal capable of discriminating a particular target in various environments.

After introducing our propagation and target models in section II, a finite dimensional approximation is obtained in section III. The criterion for signal selection and the sequential test procedure are discussed in section 2.3. Section 2.4 presents examples of radar classification models in which we apply our approach. In the remainder of the chapter we give simulation results that show the improvements that can be obtained by applying the procedure.

2.2 Target Model and Problem Formulation

Consider a point target at range r and radial velocity v relative to a transmitter and receiver located at the same position as shown in Fig. 2.1. Suppose an arbitrary signal $s(t)$ is transmitted. The received signal
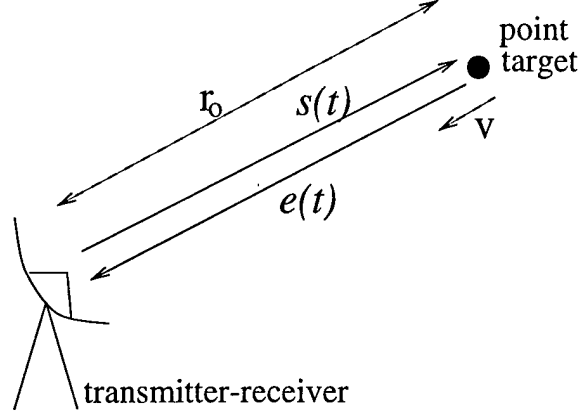
Figure 2.1: Source-Target setup

after reflection from a point target is given by [17]

$$r(t) = As(y(t - x)) \tag{2.2.1}$$

where

$$\begin{aligned} x &= \frac{2r}{c}, \\ y &= \frac{c - v}{c + v}. \end{aligned} \tag{2.2.2}$$

Here, A is a constant that depends on the range of the object, its reflectivity properties, frequency of operation among other factors and c is the speed of propagation. We will assume that we operate in a high frequency range where the reflectivity remains constant over a wide range of frequencies. The return signal is therefore a delayed and time-scaled version of the transmitted signal. Now suppose that we have a target that is composed of a continuum of point targets each having a return given by (2.2.1). It is then clear by the principle of superposition that the echo from such a target is given by

$$r(t) = \int_0^\infty dy \int_0^\infty dx \, D_c(x, y) s(y(t - x)). \tag{2.2.3}$$

Here, $D_c(x, y)$ is the target *reflectivity function* which describes the reflective properties of a point target at range $cx/2$ and radial velocity $(1 - y)c/(1 + y)$ similar to the constant A in (2.2.1). Note that the limits of the integration are determined by the physics of the underlying problem. Since x is related to the distance by (2.2.2), it takes only positive values. Similarly, the scale parameter y is related to the radial velocity v by (2.2.2) where $v < c$.

The model described in (2.2.3) is based on several inherent assumptions.

1. Objects are modeled as noninteracting point scatterers.

2. Each point scatterer is frequency-independent.

3. Operation is at a high enough frequency which allows the use of the geometric optics approximation to Maxwell's equations. Under this approximation propagation and reflection are very similar to their optical version.

4. Measurement is free from clutter or measurement noise.

The latter assumption will be dropped in section 2.3 where we will assume additive noise as well as clutter added to our observation model. The same model given by (2.2.3) has recently been used to derive algorithms for wideband radar imaging [2], [18], [19].

5

2.2.A Problem Formulation

We can now formulate the classification problem that we study in this chapter. We would like to determine a set of signals $s_n(t)$, $n = 1, 2, \ldots$, each of finite duration T_0, to minimize the expected decision time in classifying the radar targets of interest subject to achieving a desired classification performance (probability of error). A decision is made by measuring the return given by

$$r_n(t) = \int_0^\infty dy \int_0^\infty dx \, D_c(x,y) s_n(y(t-x)) + w(t) \tag{2.2.4}$$

where $w(t)$ is a noise process.

Instead of solving the problem in its present continuous form, we will consider a discrete equivalent which we derive in the next section. We will then derive a suboptimal solution to that problem in section 2.3.

2.2.B Finite Dimensional Approximation

In this section introduce a discretization of the Fourier transform of (2.2.3) with respect to t. This approximation is presented for two reasons. First, it allows us to simulate our results for arbitrary reflectivity functions. Second, it leads to a simpler presentation of the signal selection problem.

We will assume that $D_c(x,y)$ has finite L_1 norm with respect to x, that $D_c(x,y)$ has finite energy, i.e.,

$$\int_0^\infty dx \int_0^\infty dy \, |D_c(x,y)|^2 < \infty. \tag{2.2.5}$$

By taking the Fourier transform of (2.2.3) we obtain

$$R(\Omega) = \int_0^\infty dy \; \Delta_c(\Omega, y) \frac{S(\frac{\Omega}{y})}{y} \tag{2.2.6}$$

where Ω is the continuous frequency variable. In the above equation $\Delta_c(\Omega, y)$ is the Fourier transform of $D_c(x,y)$ with respect to x, i.e.,

$$\Delta_c(\Omega, y) = \int_{-\infty}^\infty D_c(x,y) e^{-j\Omega x} dx. \tag{2.2.7}$$

Define $R^+(\Omega) = R(\Omega)$ for $\Omega \geq 0$ and $R^+(\Omega) = 0$ otherwise. Similarly, let $R^-(\Omega) = R(\Omega)$ for $\Omega < 0$ and $R^-(\Omega) = 0$ otherwise. By the transformation $u = \Omega/y$ in (2.2.6), we obtain for $R^+(\Omega)$

$$R^+(\Omega) = \int_0^\infty T_+(\Omega, u) S^+(u) \frac{du}{u} \tag{2.2.8}$$

where

$$T_+(\Omega, u) = \begin{cases} \Delta_c(\Omega, \frac{\Omega}{u}) & \Omega \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2.9}$$

Similarly,

$$R^-(\Omega) = \int_{-\infty}^0 T_-(\Omega, u) S^-(u) \frac{du}{|u|} \tag{2.2.10}$$

where

$$T_-(\Omega, u) = \begin{cases} \Delta_c(\Omega, \frac{\Omega}{u}) & \Omega < 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2.11}$$

If $\int |S(\Omega)|^2 \frac{d\Omega}{\Omega} < \infty$, it follows from (2.2.8) that $T_+(\Omega, u)$ is the kernel of a map $\mathbf{T}_+$ from $L_2(R_+, du/u)$ into $L_2(R_+, d\Omega)$. Next, we discretize (2.2.8) by determining a discrete approximation for $T_+(\Omega, u)$, $S^+(u)$ and $R^+(\Omega)$.

In practice $D_c(x,y)$ is given as samples on a two-dimensional rectangular grid and is defined by

$$D(m,n) = D_c(x,y)\Big|_{x=m\Delta x, y=n\Delta y} \tag{2.2.12}$$

where m and n are index variables into the grid and Δx and Δy are the sampling intervals of the variables x and y, respectively. For most targets of interest one can assume that $\Delta_c(\Omega, y)$ is essentially bandlimitted in the frequency variable Ω with a maximum frequency of Ω_{max}. To minimize aliasing we choose Δx to satisfy the Nyquist sampling rate

$$\Delta x \leq \frac{1}{2\Omega_{max}}. \tag{2.2.13}$$

With this choice of Δx, we obtain

$$\Delta(\omega, y) \approx \Delta_c(\Omega \Delta x, y) \tag{2.2.14}$$

where $\omega = \Omega \Delta x$ is the discrete frequency variable.

To derive the sampling interval Δy, we note that $D_c(x, y)$ has finite support and therefore

$$y_{min} \leq y \leq y_{max}. \tag{2.2.15}$$

Since we are seeking a finite-dimensional approximation to the kernel $\mathbf{T}_+$, sampling in the y direction must be at a sufficiently high rate to allow for the recovery of $T_+(\Omega, u)$ in the u direction for each fixed value of Ω. The functional relationship between u and y for a fixed Ω is

$$u = \frac{\Omega}{y} \tag{2.2.16}$$

which for evenly spaced sampling points in the y domain produces a nonuniform sampling grid in u. For small values of u we have finer sampling while for larger values we have coarser sampling. Since y is bounded by y_{min} and y_{max}, u is also bounded by

$$\frac{\Omega}{y_{max}} \leq u \leq \frac{\Omega}{y_{min}}. \tag{2.2.17}$$

Soumekh [20] described several reconstruction schemes when the data is defined on an evenly sampled grid in one domain and it is required to obtain the data in another domain that has a one-to-one relationship to the first. He also determines the requirements on the sampling frequency and concludes that a signal can be recovered from its unevenly spaced samples provided that the variable sampling rate satisfy the Nyquist criterion, i.e.,

$$u_{n+1} - u_n < \frac{1}{2T_0} \qquad \text{for all } n. \tag{2.2.18}$$

Here T_0 is the support of the transmitted signals and $u_n = \Omega/y_n$. The sampling interval Δy can then be obtained by substituting for T_0 in (2.2.18).

Having determined the sampling intervals for both variables x and y, we can now write the discrete approximation of equation (2.2.6) for positive frequencies as follows

$$R^+(k) = \Delta y \sum_n \frac{\Delta_+(k,n)}{n\Delta y} S^+(\frac{k}{n\Delta y}). \tag{2.2.19}$$

where k is the discrete frequency index. Note that this equation cannot be written in a matrix vector form since the signal vector changes for different values of k. By reordering the matrix $[\Delta_+(k,n)]$, a matrix-vector form for this equation can be obtained. The resulting matrix is the discrete scaled version of $\mathbf{T}_+$.

In Appendix 2.A, we show how (2.2.8) can be written in the matrix-vector form

$$\mathbf{r} = \mathbf{As} \tag{2.2.20}$$

where the entries of $\mathbf{A}$ correspond to the interpolated samples of $[\Delta_+(k,n)]/n\Delta y$, vector $\mathbf{r}$ has entries corresponding to the samples of $R^+(k)$ and vector $\mathbf{s}$ has entries corresponding to rearranged and interpolated samples of $S^+(\frac{k}{n\Delta y})$.

Note that when the target is observed in clutter, which has its own reflectivity function, the discretization of the reflectivity function of this clutter is added to the matrix $\mathbf{A}$ in (2.2.20). The matrix $\mathbf{A}$, therefore, contains in general both target and clutter information.

2.3 General Signal Selection Strategy

The purpose of a radar classifier is to measure the backscatter from an unknown target and ascertain which of the possible targets can be associated with the return. We shall confine ourselves to two-class problems. A class consists of a collection of targets whose reflectivity function or matrix share common properties. As an example consider the backscatter from fighter jets versus civilian airplanes. Due to the larger size of a civilian aircraft the resulting backscatter will be spread in delay and scale more than the backscatter from a fighter jet. Since the backscatter is related to the reflectivity function through (2.2.3), we can form two classes each with a representative reflectivity function. The representative reflectivity function of a civilian plane will generally have a wider spread in delay and scale than a fighter jet. We shall formulate our binary classification problem as a test of statistical hypotheses.

2.3.A The Kullback-Leibler Information Numbers

Let the null hypothesis and the alternative hypothesis each denote one of the target classes. We shall use the finite-dimensional observation model developed in the previous section. In this model, our representative class reflectivity matrices and clutter will be denoted by $\mathbf{A}_0$ and $\mathbf{A}_1$. At time i, our observation $\mathbf{r}_i$ is an M-component vector. In general, our test is performed repeatedly with possibly different signals each time. The received signal under both hypotheses is therefore given by

$$H_0: \quad \mathbf{r}_i = \mathbf{A}_0 \mathbf{s}_i + \mathbf{n}_i \qquad i = 1, 2, \ldots$$
$$H_1: \quad \mathbf{r}_i = \mathbf{A}_1 \mathbf{s}_i + \mathbf{n}_i \qquad i = 1, 2, \ldots$$

where $\mathbf{n}_i$ is the observation noise.

Recall that our goal is to devise a sequential test procedure for this problem by selecting a sequence of signals $\mathbf{s}_i, i = 1, 2, \ldots$. After each return is received a decision is made to either stop or to continue transmission. If we stop, we accept one of the hypotheses. If we decide to continue, we choose a new signal from a set of predetermined signals. We continue this process until the observed data allow clear distinction between the hypotheses.

In principle, this sequential test should minimize the probability of incorrect decision. However, since this measure is in general difficult to compute, a discrimination measure is required that is easier to evaluate and manipulate. By optimizing this dissimilarity measure we hope to minimize the error probability. Chernoff [16] suggested an experiment design procedure that minimizes the probability of error. Under the condition that the observations are independent and identically distributed (*i.i.d.*), the procedure is asymptotically optimum,i.e., for a sufficiently large sample size, the experiment selection procedure leads to decisions that minimize the probability of error.

For our purposes, the choice of an experiment corresponds to a selection of a signal. In this procedure, after each new observation $\mathbf{r}_i$ is obtained, the more likely hypothesis is estimated based on all previous observations. Suppose we are at decision stage k and that we have received $\mathbf{r}_k$. Let

$$\vec{r}_k = [\mathbf{r}_1^T \ \mathbf{r}_2^T \ldots \mathbf{r}_k^T]^T \tag{2.3.21}$$

be the concatenation of all received signals up to stage k. The probability density function of $\vec{r}_k$ is $f_0(\vec{r}_k)$ under H_0 and $f_1(\vec{r}_k)$ under H_1. In a sequential test the likelihood ratio of $\vec{r}_k$

$$\Lambda(\vec{r}_k) = \frac{f_1(\vec{r}_k)}{f_0(\vec{r}_k)},$$

is computed at each decision stage. Computing the likelihood ratio at each decision stage can be done recursively [21]. For the given misclassification or error probabilities $P[\text{decide} H_0 | H_1]$ and $P[\text{decide} H_1 | H_0]$, we compare the likelihood ratio to the thresholds $A > 1 > B$. Specifically, we select H_1 if $\Lambda(\vec{r}_k) > A$ and H_0 if $\Lambda(\vec{r}_k) < B$. The thresholds A and B are derived from the desired error probabilities [22]. As $A \to \infty$ and $B \to 0$, the error probabilities $P[\text{decide} H_0 | H_1]$ and $P[\text{decide} H_1 | H_0] \to 0$ at the expense of increasing the decision time. For our signal selection problem, we need to find signals that advance the likelihood ratio towards the correct boundary in the least possible time.

Chernoff proposed a solution to this problem. His approach can be summarized as follows:

$$\text{select } H_1 \text{ if } \Lambda(\vec{r}_k) > A$$
$$\text{select } H_0 \text{ if } \Lambda(\vec{r}_k) < B \tag{2.3.22}$$

otherwise determine the more likely hypothesis. If $\Lambda(\vec{r}) \geq 1$ then select the signal s_{k+1} that maximizes $KLIN_1$, otherwise select s_{k+1} that maximizes $KLIN_0$.

The distance measures $KLIN_0$ and $KLIN_1$ are defined by

$$
\begin{aligned}
KLIN_0 &= \int log(\frac{f_0(\vec{r}_k)}{f_1(\vec{r}_k)}) f_0(\vec{r}_k) d\vec{r}_k \\
&= -E[\Lambda(\vec{r}_k)|H_0] \qquad \text{and} \tag{2.3.23} \\
KLIN_1 &= E[\Lambda(\vec{r}_k)|H_1]. \tag{2.3.24}
\end{aligned}
$$

They are known as the *Kullback-Leibler information numbers* (KLIN) and in general we have $KLIN_0 \neq KLIN_1$.

The above procedure is also a *sequential probability ratio test* (SPRT). Chernoff shows that under the assumption of independent and identically distributed observations, the expected number of observations needed under a specific hypothesis to achieve a desired probability of misclassification is inversely proportional to the $KLIN$ corresponding to that hypothesis. Therefore, by maximizing the $KLIN$ after each new observation is received, we effectively minimize the average number observations needed to reach a decision. Although the observations in our case may not be independent, this result is generally true for correlated Gaussian observations as was pointed out in [23] where correlation between observations increased the decision time compared to the *i.i.d.* case. This is consistent with the fact that little information is supplied by each additional observation if they are correlated.

Chernoff's procedure has two possible limitations:

1. The procedure is essentially an SPRT where the number of observations necessary to reach a decision is random. If the two hypotheses are not sufficiently separated, which happens when the reflectivity functions differ only slightly across classes, the procedure might take a long time before reaching a decision. This can be unacceptable. To circumvent this problem Wald [22] suggested to truncate the test at a certain maximum number of observation at the cost of slight loss in performance.

2. A poor initial estimate of the more likely hypothesis might prolong reaching a decision.

In the cases considered in this chapter, the second limitation will not be a problem since the signal sets we obtain are independent of which hypothesis is more likely.

The sequence $log \Lambda_k(\vec{r}_k)$, $k = 0, 1, 2, \ldots$ constitutes a random process. A typical realization is shown in Fig. 2.2 where the two horizontal lines indicate the thresholds at which decisions in favor of either hypotheses are made when crossed by the log-likelihood ratio. At each stage where we decide to continue, the transmitted signal is selected such that the log-likelihood crosses either boundary in the least average time. This is depicted in the same figure, we also show another realization where by a better choice of signals a decision in favor of H_1 is reached in a fewer number of transmissions. The motivation for the sequential test described above is to transmit signals only when necessary to make the decision. In some instances the observations point clearly to one particular hypothesis. In these cases, only a few number of transmissions are necessary. This is contrary to the fixed sample size test where the number of observations is always the same. In the next section we apply the above procedure to two problems that might arise in radar.

2.4 Signal Selection for Gaussian Reflectivity Functions with Different Means and Same Covariance

Let us consider in some detail two models for the target and its environment. In both models we will assume that the reflectivity matrices under both hypotheses are composed of two parts: a known mean matrix and
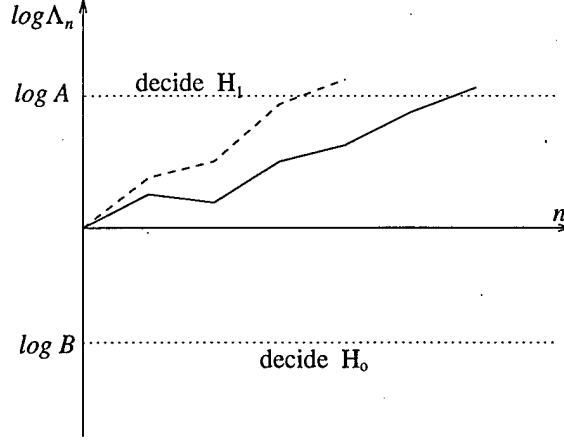
Figure 2.2: Realization of a sequential likelihood ratio test. The dashed curve shows a better choice of signals for the same test conditions.

a matrix with random entries. The random matrix is introduced for two reasons. First, it accounts for any target surface roughness that can be modeled by random variables distributed in the delay-scale plane. Since a large number of surface irregularities contributes energy in each delay-scale cell, we are led to assume by the central limit theorem that the random entries into the matrix due to surface roughness can be modeled as independent Gaussian random variables with variance σ_{sr}^2. Second, the random matrix accounts also for the unknown clutter which is a distributed interferer possibly overlapping with the delay-scale extent of the target [1]. Again, since the contribution to each radar delay-scale cell comes from a multitude of unknown reflectors, we assume that the randomness in each cell due to clutter is additive, statistically independent and Gaussian with equal variance σ_c^2. We will also assume that the randomness due to surface roughness and clutter are statistically independent.

2.4.A   Case I: Fixed Target Environment

In our first case we will assume that the target and clutter remain unchanged during the entire period of illumination. In (2.3.21), let

$$\begin{aligned} \mathbf{A}_0 &= \bar{\mathbf{A}}_0 + \mathbf{W}, \\ \mathbf{A}_1 &= \bar{\mathbf{A}}_1 + \mathbf{W}. \end{aligned} \tag{2.4.25}$$

Here, each reflectivity matrix is composed of a constant and a random part denoted by $\mathbf{W}$ which has Gaussian uncorrelated zero-mean elements with variance $\sigma_w^2 = \sigma_{sr}^2 + \sigma_c^2$. Note that the entries of $\mathbf{W}$ may occur outside the delay scale extent of the target accounting for clutter at these location in the delay-scale plane. The matrix $\mathbf{W}$ remains fixed during the illumination period. The received signal under both hypotheses is given by

$$\begin{aligned} H_0 &: \quad \mathbf{r}_i = (\bar{\mathbf{A}}_0 + \mathbf{W})\,\mathbf{s}_i + \mathbf{v}_i && i = 1, 2, \cdots \\ H_1 &: \quad \mathbf{r}_i = (\bar{\mathbf{A}}_1 + \mathbf{W})\,\mathbf{s}_i + \mathbf{v}_i. && i = 1, 2, \cdots. \end{aligned} \tag{2.4.26}$$

After each transmission, the vector $\mathbf{r}_i$ is received corrupted by a Gaussian vector $\mathbf{v_i}$ with covariance matrix $\sum_v$. This vector captures both the receiver and atmospheric noise. We will assume that $\sum_v = \sigma_v^2 \mathbf{I}$ where $\mathbf{I}$ is the identity matrix. The transmitted signals $\mathbf{s}_i$, $i = 1, 2, \ldots$ are assumed to be of unit energy.

Since a hypothesis testing problem is invariant to a bias adding transformation, we can rewrite (2.4.26) as

$$\begin{aligned} H_0 &: \quad \mathbf{r}_i = (\bar{\mathbf{A}}_0 - \bar{\mathbf{A}}_1 + \mathbf{W})\,\mathbf{s}_i + \mathbf{v}_i && i = 1, \cdots \\ H_1 &: \quad \mathbf{r}_i = \mathbf{W}\,\mathbf{s}_i + \mathbf{v}_i. && i = 1, \cdots \end{aligned}$$

10

Following the steps of the procedure outlined in the previous section, we need to determine the *KLIN* under both hypotheses. Let n be the current decision stage and let

$$\vec{r}_n = \begin{bmatrix} \mathbf{r}_1^T & \mathbf{r}_2^T & \cdots & \mathbf{r}_n^T \end{bmatrix}^T \tag{2.4.27}$$

be the vector composed of all n observations. Under the null hypothesis, $\vec{r}_n$ is a Gaussian random vector of length $(n \times L)$ with mean

$$E[\vec{r}_n | H_0] = [\mathbf{I}_n \otimes (\bar{\mathbf{A}}_0 - \bar{\mathbf{A}}_1)] \cdot \mathbf{s} = \mathbf{G} \cdot \vec{s} \tag{2.4.28}$$

where $\otimes$ denotes the Kronecker product and $\vec{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_n^T]^T$. The covariance matrix of $\vec{r}_n$ is given by

$$\mathbf{K}_r = \begin{bmatrix} \sigma_w^2 + \sigma_v^2 & \sigma_w^2 \cos\theta_{12} & \cdots & \sigma_w^2 \cos\theta_{1n} \\ \sigma_w^2 \cos\theta_{21} & \sigma_w^2 + \sigma_v^2 & \cdots & \sigma_w^2 \cos\theta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_w^2 \cos\theta_{n1} & \sigma_w^2 \cos\theta_{n2} & \cdots & \sigma_w^2 + \sigma_v^2 \end{bmatrix} \otimes \mathbf{I}_M \tag{2.4.29}$$

where $\cos\theta_{ij} = \mathbf{s}_i^T \mathbf{s}_j$ and M is the row dimension of $\bar{\mathbf{A}}_0$ or $\bar{\mathbf{A}}_1$. Under the alternative hypotheses, the mean and covariance of $\vec{r}_n$ are $E[\vec{r}_n | H_1] = \mathbf{0}$ and $\mathbf{K}_r$, respectively.

At each decision stage n the observation vector $\vec{r}_{n-1}$ is appended with a new observation $\mathbf{r}_n$ and a new log-likelihood ratio is computed as follows:

$$\log \Lambda(\vec{r}_n) = -\frac{1}{2}\vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \vec{s} + \vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \vec{r}_n. \tag{2.4.30}$$

To determine the *KLIN* number under H_0 we compute

$$\begin{aligned} KLIN_0 &= -E[\log \Lambda(\vec{r}_n) | H_0] \\ &= \frac{1}{2}\vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \vec{s}. \end{aligned} \tag{2.4.31}$$

Similarly, the *KLIN* under H_1 is given by

$$\begin{aligned} KLIN_1 &= E[\log \Lambda(\vec{r}_n) | H_1] \\ &= \frac{1}{2}\vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \vec{s}. \end{aligned} \tag{2.4.32}$$

According to the optimal procedure in (2.3.22), we need to determine the more likely hypothesis in order obtain the *KLIN* to maximize. From (2.4.31) and (2.4.32) it is clear that under both hypotheses these measures are the same. Therefore, the same signal set that maximizes $KLIN_0$ at each stage also maximizes $KLIN_1$. This has the implication that at each decision stage we do not need to determine which hypothesis is more likely to be true and can immediately send out the next optimal signal if a terminal decision cannot be made. This overcomes one of the shortcomings in Chernoff's procedure as was pointed out in the previous section. It also results in a fixed waveform selection strategy.

2.4.A.1 *Maximizing the KLIN*

We now study the maximization of the *KLIN* given in (2.4.31) or (2.4.32) under various conditions for the ratio σ_v^2/σ_w^2. We shall use $KLIN^{(n)}$ to refer to either Kullback-Leibler numbers at the nth decision stage. We shall also drop the $\frac{1}{2}$ factor and refer to *KLIN* and $\vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \vec{s}$ interchangeably.

First, we determine the initial signal to transmit. This signal is sent at stage $n = 1$ when no prior observation is available. This signal is found by maximizing

$$\vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \vec{s} = \frac{\mathbf{s}_1^T \bar{\mathbf{A}}^T \bar{\mathbf{A}} \mathbf{s}_1}{\sigma_w^2 + \sigma_v^2} \tag{2.4.33}$$

where $\bar{\mathbf{A}} = \bar{\mathbf{A}}_0 - \bar{\mathbf{A}}_1$. The right hand side followed from the definition of $\mathbf{K}_r$ in (2.4.29). The maximizing signal $\mathbf{s}_1^*$ is clearly the eigenvector of $\bar{\mathbf{A}}^T \bar{\mathbf{A}}$ corresponding to its maximum eigenvalue λ_1.

Suppose that we are at the nth decision stage in our procedure (2.3.22) and that a terminal condition was not reached. We need to determine the next signal, $\mathbf{s}_n^*$, to transmit next by maximizing the $KLIN$. This signal can be obtained as follows:

$$\mathbf{s}_n^* = \arg\left[\max_{\|\mathbf{s}_n\|=1} \vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \vec{s}\right] = \arg\left[\max_{\|\mathbf{s}_n\|=1} \left[\begin{array}{ccccc} \mathbf{s}_1^* & \mathbf{s}_2^* & \cdots & \mathbf{s}_{n-1}^* & \mathbf{s}_n \end{array}\right]^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G} \begin{bmatrix} \mathbf{s}_1^* \\ \mathbf{s}_2^* \\ \ldots \\ \mathbf{s}_{n-1}^* \\ \mathbf{s}_n \end{bmatrix}\right] \quad (2.4.34)$$

where the signals $\mathbf{s}_1^*, \mathbf{s}_2^*, \cdots, \mathbf{s}_{n-1}^*$ are the optimum signals transmitted up to stage $n-1$. We consider two cases.

1. **Noiseless Case**

   In this case the observation noise variance is assumed to be zero. Substituting $\sigma_v^2 = 0$ into $\mathbf{K}_r^{-1}$ in (2.4.34), we find that $\mathbf{s}_n^*$ is the eigenvector of $\bar{\mathbf{A}}^T\bar{\mathbf{A}}$ corresponding to the nth largest eigenvalue. Therefore, in the noiseless case maximum discrimination is provided by transmitting the eigenvectors of $\bar{\mathbf{A}}^T\bar{\mathbf{A}}$. The sequence of transmission is according to a decreasing order of magnitude of the eigenvalues. The details of this optimization problem is given in Appendix 2.B. The $KLIN$ after n transmissions is given by

$$KLIN^{(n)} = \frac{1}{\sigma_w^2} \sum_{k=1}^{n} \lambda_k. \quad (2.4.35)$$

2. **Noisy Case**

   When $\sigma_v^2 > 0$ and we are at the nth decision stage, all signals up stage $n-1$ are either orthogonal and selected as in the noiseless case or some of them were repeatedly transmitted. A repeated transmission improves a previous measurement by averaging the effect of noise while an orthogonal signal obtains new independent information from the observed target. We will treat the two cases separately.

   (a) *Orthogonal Transmissions up to stage $n-1$*

   In this case one can write the $KLIN$ at stage $n-1$ as

$$KLIN^{(n-1)} = \sum_{i=1}^{n-1} \frac{\lambda_i}{\sigma_v^2 + \sigma_w^2}. \quad (2.4.36)$$

   We shall assume that the eigenvalues of $\bar{\mathbf{A}}^T\bar{\mathbf{A}}$ are ordered such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1}$. At stage n, we can either transmit a new signal corresponding to λ_n and add $\lambda_n/(\sigma_v^2 + \sigma_w^2)$ to $KLIN^{(n-1)}$ or retransmit one of the signals $\mathbf{s}_i^*$, $i = 1, 2, \ldots n-1$. The selection here is determined by the value of the ratio σ_v^2/σ_w^2. If the observation noise dominates our measurements, corresponding to a larger value of σ_v^2/σ_w^2, then we can increase our discrimination by repeating a transmission so that the combined return from that transmission is more accurate due to averaging. If, on the other hand, the clutter dominates our measurement, a repeated transmission would not add any new discrimination information since the clutter is assumed to remain unchanged during illumination. In this case, a new independent transmission provides an increase in the discrimination information. In Appendix 2.B, we show that a retransmission of $\mathbf{s}_i^*$, for example, causes the corresponding term in the $KLIN^{(n-1)}$ in (2.4.36) to change to

$$\frac{\lambda_i}{\sigma_v^2 + \sigma_w^2/2} \quad (2.4.37)$$

   thereby effectively reducing the effect of the observation noise on this particular measurement. If we decide to retransmit a signal it is simple to see that it will be $\mathbf{s}_1^*$ since $\lambda_1 \geq \lambda_i$, $i \neq 1$. In Appendix 2.B we show that retransmission occurs if the nth eigenvalue, λ_n, of $\bar{\mathbf{A}}^T\bar{\mathbf{A}}$ is such that

$$\frac{\sigma_v^2}{\sigma_w^2} > 2\left(\frac{\lambda_n}{\lambda_1}\right)/\left(1 - \frac{\lambda_n}{\lambda_1}\right) \quad (2.4.38)$$

12

then retransmitting the first eigenvector of $\bar{\mathbf{A}}^T \bar{\mathbf{A}}$, $\mathbf{s}_1^*$, results in a larger discrimination than transmitting the nth eigenvector.

(b) *Some Repeated Transmissions*

In this case we assume that up to stage $n-1$ we have transmitted the signals $\mathbf{s}_1^*, \mathbf{s}_2^*, \ldots, \mathbf{s}_j^*$, $n_1, n_2, \ldots, n_j$ times, respectively where

$$\sum_{i=1}^{j} n_i = n - 1. \tag{2.4.39}$$

We need to determine the waveform to transmit at the nth stage. This is best done by rewriting the *KLIN* at stage n as

$$KLIN^{(n)} = KLIN^{(n-1)} + \Delta KLIN \tag{2.4.40}$$

where $KLIN^{(n-1)}$ is the discrimination information obtained from the signals transmitted up to stage $n-1$ and $\Delta KLIN$ is the *discrimination gain* due to the new signal $\mathbf{s}_n$. To determine $\mathbf{s}_n^*$, we maximize the discrimination gain $\Delta KLIN$. The choice of whether to retransmit a previous signal or transmit a new one follows the same reasoning as in the previous case. In Appendix 2.B we perform the above maximization and obtain $\mathbf{s}_n^*$. This waveform can either be a retransmission of a previous eigenvector or the new eigenvector corresponding to λ_{j+1}. If the condition

$$\frac{\lambda_{j+1}}{1 + \frac{\sigma_v^2}{\sigma_w^2}} < \max_i \frac{\lambda_i'(\frac{\sigma_v^2/n_i}{\sigma_w^2})}{(n_i + 1) + \frac{\sigma_v^2}{\sigma_w^2}} \tag{2.4.41}$$

is satisfied then we retransmit the kth eigenvector of $\bar{\mathbf{A}}^T \bar{\mathbf{A}}$ where

$$k = arg[\max_i \frac{\lambda_i'(\frac{\sigma_v^2/n_i}{\sigma_w^2})}{(n_i + 1) + \frac{\sigma_v^2}{\sigma_w^2}}] \tag{2.4.42}$$

and

$$\lambda_i' = \frac{\lambda_i}{(1 + \frac{\sigma_v^2/n_i}{\sigma_w^2})}.$$

Otherwise, we transmit the new eigenvector. Note that the previous case is just a special case with $n_1 = n_2 = \cdots = n_{n-1} = 1$ and the condition in (2.4.41) reduces to (2.4.38). As in the previous case, we have the choice of either transmitting a signal along a new direction or retransmit a previous signal. For given targets, the choice is made based on the value of the ratio σ_v^2/σ_w^2.

2.4.B Case II: Variable Target Environment

In the previous example we assumed that the target and clutter remain unchanged during illumination. As a result, we fixed the random matrix $\mathbf{W}$ in (2.4.26). In this section we assume that the received signal under both hypotheses is given by

$$\begin{aligned} H_0: \quad & \mathbf{r}_i = (\bar{\mathbf{A}}_0 + \mathbf{W}_i)\,\mathbf{s}_i + \mathbf{v}_i & i = 1, 2, \cdots \\ H_1: \quad & \mathbf{r}_i = (\bar{\mathbf{A}}_1 + \mathbf{W}_i)\,\mathbf{s}_i + \mathbf{v}_i. & i = 1, 2, \cdots \end{aligned} \tag{2.4.43}$$

As with our previous model, we assume that each reflectivity matrix is composed of a constant and a random part denoted by $\mathbf{W}_i$ which Gaussian uncorrelated zero-mean elements with variance $\sigma_{w_i}^2 = \sigma_{sr_i}^2 + \sigma_{c_i}^2$. The matrix $\mathbf{W}_i$ changes after each transmission and we will assume that each entry w_{kj} is statistically independent from the entries in the same kj location of some other random matrix $\mathbf{W}_j$, $j \neq i$. We also assume that under both hypotheses the random matrix $\mathbf{W}_i$ has the same statistical characteristics. The transmitted signals $\mathbf{s}_i$, $i = 1, 2, \ldots$ are again assumed to be of unit energy. After each transmission, the vector $\mathbf{r}_i$ is received corrupted by a Gaussian vector $\mathbf{v}$ with covariance matrix $\Sigma_v = \sigma_v^2 \mathbf{I}$.

The analysis of this model proceeds along the same lines as the previous model and we shall only outline it and mention the relevant results. Our observation vector at the nth decision stage is

$$\vec{r}_n = \begin{bmatrix} \mathbf{r}_1^T & \mathbf{r}_2^T & \cdots & \mathbf{r}_n^T \end{bmatrix}^T. \tag{2.4.44}$$

Its mean and covariance are given by

$$E[\vec{r}_n|H_0] = [\mathbf{I}_n \otimes (\bar{\mathbf{A}}_0 - \bar{\mathbf{A}}_1)] \cdot \mathbf{s} = \mathbf{G} \cdot \vec{s} \tag{2.4.45}$$

$$E[\vec{r}_n|H_1] = \mathbf{0} \qquad \text{and} \tag{2.4.46}$$

$$\mathbf{K}_{r_0} = \mathbf{K}_{r_1} = \text{diag}\{\sigma_{w_1}^2 + \sigma_v^2, \sigma_{w_2}^2 + \sigma_v^2, \ldots, \sigma_{w_n}^2 + \sigma_v^2\} \otimes \mathbf{I}_M \tag{2.4.47}$$

where M is the row dimension of $\bar{\mathbf{A}}_0$ or $\bar{\mathbf{A}}_1$.

It is simple to see that for this case also the Kullback-Leibler numbers are equal. This results in a fixed waveform selection strategy where one signal set is sufficient for our test procedure. The $KLIN$ for this case is given by

$$KLIN = \frac{1}{2}\mathbf{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G} \mathbf{s} \tag{2.4.48}$$

$$= \sum_{i=1}^{n} \frac{\mathbf{s}_i^T \bar{\mathbf{A}}^T \bar{\mathbf{A}} \mathbf{s}_i}{\sigma_{w_i}^2 + \sigma_v^2} \tag{2.4.49}$$

where $\bar{\mathbf{A}} = \bar{\mathbf{A}}_0 - \bar{\mathbf{A}}_1$. Maximizing the $KLIN$ for this case is much simpler than in the previous example since the observations are always independent. This is apparent from the diagonal covariance matrix $\mathbf{K}_r$. At stage n the $KLIN$ is maximized by

$$\mathbf{s}_n^* = \arg\left[\max_{\|\mathbf{s}_n\|=1} \mathbf{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G} \mathbf{s}\right] \tag{2.4.50}$$

$$= \mathbf{s}_1^* \tag{2.4.51}$$

which is the eigenvector of $\bar{\mathbf{A}}^T \bar{\mathbf{A}}$ corresponding to the largest eigenvalue. The $KLIN$ is therefore given by

$$KLIN = \sum_{i=1}^{n} \frac{\lambda_1}{\sigma_{w_i}^2 + \sigma_v^2}. \tag{2.4.52}$$

The result indicates that when the target is observed in a continually changing environment, repeated transmission of one carefully selected signal provides the largest discrimination information.
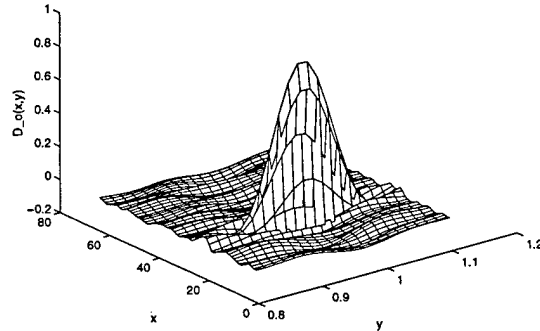

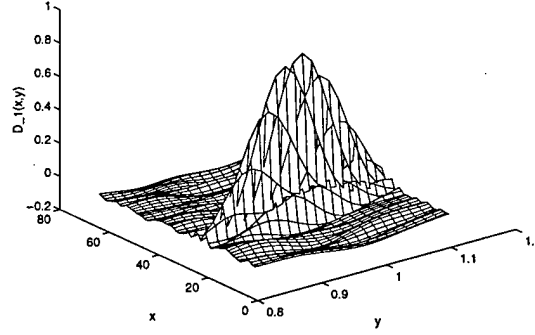Figure 2.3: Mean reflectivity function of the first target

14

Figure 2.4: Mean reflectivity function of the second target

2.5 Simulation Results

We simulated a case with 2 targets and obtained the optimal waveforms for different values of the ratio σ_v^2/σ_w^2 for the first case of section 2.4. The reflectivity functions of these targets as a function of the scale parameter y and the delay parameter x are shown in Fig. 2.3 and 2.4. We assumed that these targets representatives of a larger class of targets and determined the associated reflectivity matrices $\bar{A}_0$ and $\bar{A}_1$ as described in Chapter 3. The eigenvectors of $\bar{A}_0 - \bar{A}_1$ were determined. These are the optimal discrimination signals for a noiseless environment. They also constitute the library of waveforms from which we select our transmitted signals.

We examined the performance of the procedure we developed using the optimal signals relative to a rectangular pulse waveform defined by:

$$rect(\frac{t}{T}) = \begin{cases} 1 & 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases} \tag{2.5.53}$$

where T is small compared to the delay extent of the target. We vary the ratio σ_v^2/σ_w^2 and examine the performance by drawing the receiver operating characteristics (ROC) for a given value of n. The probability of false alarm for a given n is defined as $P_F|_n = Prob\{accept H_1|H_0, n\}$ and the probability of detection for a given n is defined as $P_D|_n = Prob\{accept H_1|H_1, n\}$. Figure 2.5 compares the performance of the optimal waveform to that of the $rect(\frac{t}{T})$ waveform in two cases. Once in a noiseless environment with $\sigma_v^2 = 0$ and the other when noise is present. For both cases, the optimal waveform performs much better than the arbitrary signal. In Fig. 2.6 noise is added such that the optimal set consists of the first two orthonormal signals for $n = 2$ and then a repetition of the first signal when $n = 3$. Obviously, the performance of the optimal waveforms is again much better than the rectangular signal. In Fig. 2.7 the observation noise variance is increased such that the optimal waveforms are repetitions of the first one. This case corresponds to $\sigma_v^2/\sigma_w^2 \gg 2(\frac{\lambda_2}{\lambda_1})/1 - (\frac{\lambda_2}{\lambda_1})$ with noise dominating the received signal. Again, clearly the optimal waveforms outperform the rectangular signal.

2.6 Conclusion

We have proposed a signal selection procedure for radar target classification. The signals were designed to maximize a dissimilarity measure between target classes. During the classification process, the signals are transmitted sequentially in an order determined by the noise and clutter powers. The classifier continues to transmit the signals until a decision can be made with a specified level of confidence. The signals are obtained before actual classification from the reflectivity functions of the targets of interest. The signals are chosen such that at actual classification they maximize the *Kullback-Leibler information numbers* as a measure of dissimilarity between the returns from different targets.
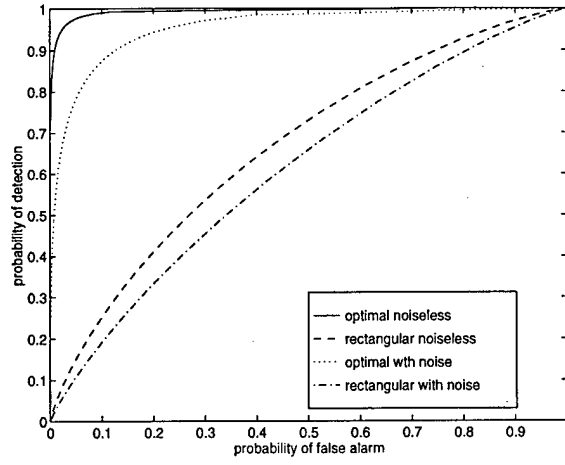
15

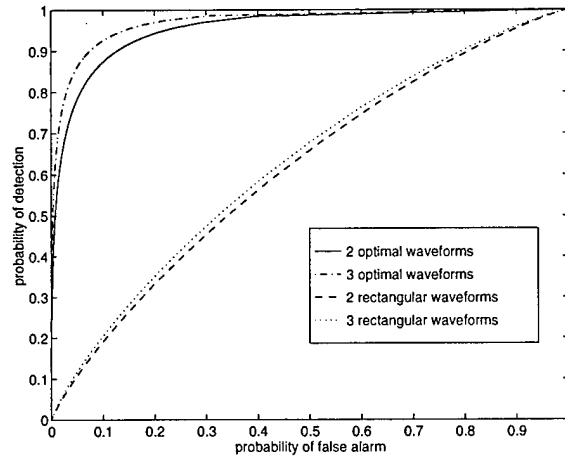Figure 2.5: ROC curves for $n = 1$ with optimal and *rect* waveforms.

Figure 2.6: ROC curves for $n = 2$ and $n = 3$ for $2(\frac{\lambda_3}{\lambda_1})/1 - (\frac{\lambda_3}{\lambda_1}) < \sigma_v^2/\sigma_w^2 < 2(\frac{\lambda_2}{\lambda_1})/1 - (\frac{\lambda_2}{\lambda_1})$.
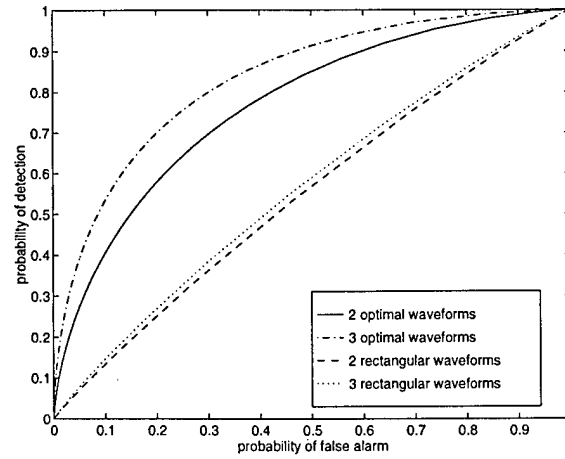

Figure 2.7: ROC curves for $n = 2$ and $n = 3$ in a high noise situation.

16

One can extend the above procedure in several directions

1. In this chapter we only considered signal selection in a two-class problem. Naturally, we would like to extend the results to the multi-class case. The extension to such cases can still employ results from the two-class case. For example, one could find the signals that maximally discriminates between the most likely target hypothesis and the hypothesis closest to it [27]. These signals can be obtained using techniques developed in this chapter.

2. Other target models than the ones described in this chapter can also be considered. We discussed the Gaussian unequal mean and equal covariance models. Other models such as Gaussian equal means and unequal covariances can also be considered. What needs to be seen is whether these models are as analytically tractable as the model we discussed in this chapter. The other issue such models can pose is the issue of estimating the more likely hypothesis during the course of classification and whether an erroneous initial estimate can still be recovered in an acceptable amount of time.
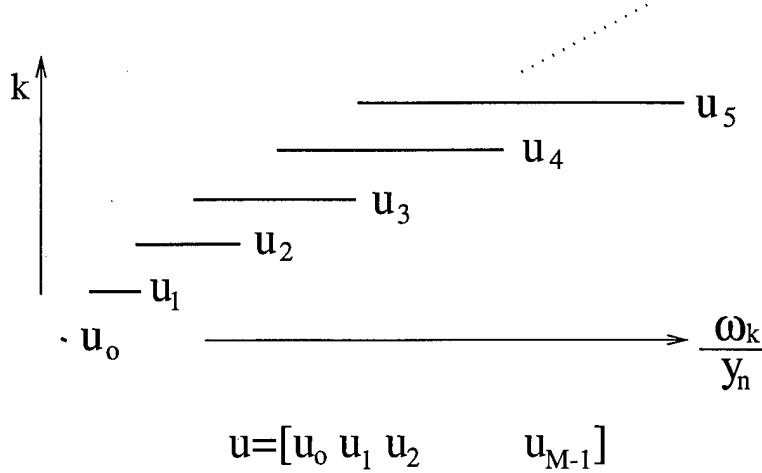
Figure 2.8: Forming the vector $\mathbf{u}$ from overlapping vectors $\mathbf{u}_k$.

Appendix 2.A: Finite-Dimensional Approximation

The matrix $\mathbf{A}$ in (2.2.20) can be obtained as follows. Let $y \in [y_1, y_J]$ and $\omega \in [0, \pi]$. For each value of $\omega_k = \pi k/M$, $k = 1, \cdots, M$ we construct the vector $\mathbf{u}_k = \omega_k [\frac{1}{y_1}, \frac{1}{y_{J-1}}, \cdots, \frac{1}{y_J}]$. The values in those vectors constitute the frequencies at which the signal vector $\mathbf{s}$ has to be determined. Figure 2.8 shows the overlap that occurs between the values of these vectors for different k.

To construct $\mathbf{s}$ and the rows of $\mathbf{A}$ at these frequency values, we form a vector $\vec{u}$ by concatenating together all $\mathbf{u}_k$ for all k. Since some of the values of $\vec{u}$ will either be redundant or occurring out of order due to the overlap, an ascending reordering of the values of $\vec{u}$ is done and repeated values are eliminated. The indexes in $\vec{u}$ where the changes occur are recorded and the matrix

$$\left[\frac{\Delta_+(\omega, y)}{y}\right] = \begin{pmatrix} \frac{\Delta_+(\omega_1, y_J)}{y_J} & \frac{\Delta_+(\omega_1, y_{J-1})}{y_{J-1}} & \cdots & \frac{\Delta_+(\omega_1, y_1)}{y_1} \\ \frac{\Delta_+(\omega_2, y_J)}{y_J} & \frac{\Delta_+(\omega_2, y_{J-1})}{y_{J-1}} & \cdots & \frac{\Delta_+(\omega_2, y_1)}{y_1} \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\Delta_+(\omega_M, y_J)}{y_J} & \frac{\Delta_+(\omega_M, y_{J-1})}{y_{J-1}} & \cdots & \frac{\Delta_+(\omega_M, y_1)}{y_1} \end{pmatrix}$$

is reordered according to these changes.

To see how these changes are applied to $\left[\frac{\Delta_+(\omega, y)}{y}\right]$, suppose that $\frac{\omega_m}{y_{j-1}} < \frac{\omega_{m+l}}{y_i} < \frac{\omega_m}{y_j}$ where $l \geq 1$ and i, j and m are some integers. Also assume that the value $\frac{\omega_{m+l}}{y_i}$ has not appeared in any vector $\mathbf{u}_k$ for $k \leq m + l$. In this case a 0 is placed temporarily in the rows k of $\left[\frac{\Delta_+(\omega, y)}{y}\right]$ for $k < m + l$ and in the column position between $j - 1$ and j, thus increasing the column dimension by one. This procedure is repeated for all M rows. Since the position of the temporary zeros in $\left[\frac{\Delta_+(\omega, y)}{y}\right]$ actually corresponds to a certain frequency of $S(\frac{\omega_k}{y})$, its value is determined from adjacent known values by interpolation.

We still need to determine the values of the rows of the matrix $\left[\frac{\Delta_+(\omega, y)}{y}\right]$ on an equispaced grid. This can be done by applying one of the reconstruction techniques described in [20] provided that the Nyquist criterion on Δy is satisfied as discussed earlier. The result is the matrix $\mathbf{A}$ in (2.2.20) and we will assume it is in general of size $M \times L$ with $L > M$. A similar procedure can be used to obtain the discrete version of $\mathbf{T}_-$.

18

Appendix 2.B : Solution of the Optimization Problem

In this appendix, we determine the nth optimal signal, SS_n^*, by carrying out the maximization in (2.4.51). The derivation of this signal will have 3 special cases. Each case discusses a particular scenario for the signals transmitted up to that stage and the corresponding ratio σ_v^2/σ_w^2. In the first case, a noiseless environment is assumed where $\sigma_v^2 = 0$. In the second case, we find s_n^* when all previously transmitted signals were orthogonal and correspond to eigenvectors of the matrix $\bar{A}^T\bar{A}$. The choice of s_n will depend on the ratio σ_v^2/σ_w^2. In the third case, we assume that multiple signal transmissions have occurred during previous decision stages.

2.B.A   Noiseless Case ($\sigma_v^2 = 0$)

We substitute $\sigma_v^2 = 0$ into K_r in (2.4.29) and use an induction argument to show that

$$\max_{\|s_n=1\|} \vec{s}^T G^T K_r^{-1} G \vec{s} = \frac{1}{\sigma_w^2} \sum_{k=1}^{n} \lambda_k. \tag{2.B.1}$$

Here, λ_k is the kth largest eigenvalue of $\bar{A}^T\bar{A}$. The kth transmitted signal is the eigenvector of $\bar{A}^T\bar{A}$ corresponding to λ_k, $k = 1, 2, \ldots, n$. Hence, in the noiseless case, the signals are the eigenvectors of $\bar{A}^T\bar{A}$ transmitted in the order of decreasing magnitude of the corresponding eigenvalues.

2.B.B   No repeated transmissions

We study this case when $\sigma_v^2 > 0$. We will assume that no repeated transmissions occurred up to stage $n-1$, where $n \geq 2$, and that the transmitted signals were the eigenvectors of $\bar{A}^T\bar{A}$ corresponding to the $n-1$ largest eigenvalues. In section 2.4 we saw that the first signal to transmit was the first eigenvector of $\bar{A}^T\bar{A}$. Therefore, assuming that the previous transmissions are the eigenvectors of $\bar{A}^T\bar{A}$ is at least valid for $n = 2$. We shall see in the sequel that this assumption is generally valid under certain conditions on the ratio σ_v^2/σ_w^2.

Let s_i^*, $i = 1, \ldots, n-1$ be the eigenvectors of $\bar{A}^T\bar{A}$ corresponding to the ith largest eigenvalue. First, we determine the $KLIN$, and need to compute K_r^{-1}. It can be shown that

$$det(K_r) = (\sigma_w^2)^n (1 + \frac{\sigma_v^2}{\sigma_w^2})^{n-2} [(1 + \frac{\sigma_v^2}{\sigma_w^2})^2 - \sum_{i-1}^{n-1} (s_n^T s_i^*)^2] \tag{2.B.2}$$

and after some algebra we get

$$\vec{s}^T G^T K_r^{-1} G \vec{s} = \frac{\frac{1}{\sigma_w^2}[\sum_{k=1}^{n-1} \lambda_k((1+\frac{\sigma_v^2}{\sigma_w^2})^2 - 2(1+\frac{\sigma_v^2}{\sigma_w^2})(s_n^T s_k^*)^2 - \sum_{\substack{j=1 \\ j \neq k}}^{n-1}(s_n^T s_j^*)^2) + (1+\frac{\sigma_v^2}{\sigma_w^2})^2 s_n^T G^T G s_n]}{(1+\frac{\sigma_v^2}{\sigma_w^2})[(1+\frac{\sigma_v^2}{\sigma_w^2})^2 - \sum_{i=1}^{n-1}(s_n^T s_i^*)^2]}. \tag{2.B.3}$$

Second, we solve for s_n^* in (2.4.51). This is a constrained optimization problem since we are restricting s_n to have unit norm.

Let $V = span\{s_1^* s_2^* \cdots s_{n-1}^*\}$ and let

$$s_n = a_1 s_1^* + a_2 s_2^* + \cdots + a_n s^\perp \tag{2.B.4}$$

where $s^\perp \in V^\perp$ and $\sum_{i=1}^{n} a_i^2 = 1$. This assumes that s_n has components along all or some of the previously transmitted signals. It will turn out later that only a_1 or a_n can have a nonzero value. Substituting (2.B.4) into (2.B.3) and simplifying, we obtain

$$\vec{s}^T G^T K_r^{-1} G \vec{s} = \frac{\frac{1}{\sigma_w^2}[x^2(\lambda + \sum_{k=1}^{n-1}\lambda_k) + a^T Q a]}{x[x^2 - a^T a]} \tag{2.B.5}$$

$$= f(a, \lambda) \tag{2.B.6}$$

19

where $\lambda = \mathbf{s}^\perp \bar{\mathbf{A}}^T \bar{\mathbf{A}} \mathbf{s}^\perp$ and will be determined, $x = (1 + \frac{\sigma_x^2}{\sigma_w^2})$, $\mathbf{a} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} \end{bmatrix}$ and $\mathbf{Q} = [q_{ii}]$ is a diagonal matrix with entries $q_{ii} = x^2(\lambda_i - \lambda) - 2x\lambda_i - \sum_{j \neq i}^{n-1} \lambda_j$. To solve this maximization problem we first maximize $f(\mathbf{a}, \lambda)$ with respect to $\mathbf{a}$ over the unit sphere $\|\mathbf{a}\|_2^2 \leq 1$ while holding λ fixed. This can be done using the Lagrange multiplier technique. Then we substitute the maximizing value of $\mathbf{a}$ into $f(\mathbf{a}, \lambda)$ and maximize with respect to λ to obtain an explicit form for $\mathbf{s}_n^*$.

The maximum of $f(\mathbf{a}, \lambda)$ with respect to $\mathbf{a}$ is determined from its stationary points that are found from the first order Karush-Kuhn-Tucker (KKT) necessary conditions [25]. The Lagrangian function for this problem is given by

$$L(\mathbf{a}) = f(\mathbf{a}, \lambda) + \mu(\mathbf{a}^T \mathbf{a} - 1). \tag{2.B.7}$$

A stationary point $\mathbf{a}^*$ satisfies the following KKT necessary conditions

$$\begin{aligned} \nabla L(\mathbf{a}^*) &= \mathbf{0} \\ &= \mathbf{H}(\mathbf{a}^*)\mathbf{a}^* + \mu \mathbf{I}\mathbf{a}^*, \\ \mu &\geq 0 \end{aligned} \tag{2.B.8}$$

where $\mathbf{H}$ is a diagonal matrix with entries

$$h_i(\mathbf{a}) = \underbrace{x^2(x^2(\lambda_i - \lambda) - 2x\lambda_i + (\lambda_i + \lambda))}_{b_i} + \sum_{j \neq i} \underbrace{(x-1)^2(\lambda_j - \lambda_i)}_{c_{ij}} a_j^2. \tag{2.B.9}$$

We shall find the stationary points of $f(\mathbf{a}, \lambda)$ in 2 cases. Once when $\mu = 0$ and then when $\mu > 0$. Each of these cases will result in one potential global maximum of $f(\mathbf{a}, \lambda)$ in $\|\mathbf{a}\| \leq 1$.

2.B.B.1 Unconstrained Stationary Points: $\mu = 0$

Equation (2.B.8) becomes

$$\mathbf{H}(\mathbf{a}) \quad \mathbf{a} = \mathbf{0} \tag{2.B.10}$$

$$\begin{bmatrix} (b_1 + \sum_{j \neq 1} a_j^2 c_{1j})a_1 \\ (b_2 + \sum_{j \neq 2} a_j^2 c_{2j})a_2 \\ \vdots \\ (b_{n-1} + \sum_{j \neq n-1} a_j^2 c_{n-1j})a_{n-1} \end{bmatrix} = \mathbf{0} \tag{2.B.11}$$

Equation (2.B.11) forms a set of $n-1$ nonlinear equations in $a_1, a_2, \cdots, a_{n-1}$. Multiplying equation i in (2.B.11) by the corresponding a_i and summing using the fact that $c_{ij} = -c_{ji}$ we obtain

$$b_1 a_1^2 + b_2 a_2^2 + \cdots + b_{n-1} a_{n-1}^2 = 0. \tag{2.B.12}$$

This equation defines a *conical surface* [26]. A conical surface is made up of straight lines going through the origin of coordinates. For simplicity and clarity we will assume that $b_i \neq 0$ for $i = 1, 2, \ldots, n-1$. Obviously, one solution to (2.B.11) is $\bar{\mathbf{a}}^1 = \mathbf{0}$. Next, we show that there are no other feasible solutions to (2.B.11). From (2.B.11) we have the following set of equations

$$\begin{aligned} (b_1 + \sum_{j \neq 1} a_j^2 c_{1j})a_1^2 &= 0 \\ (b_2 + \sum_{j \neq 2} a_j^2 c_{2j})a_2^2 &= 0 \\ &\vdots \\ (b_{n-1} + \sum_{j \neq n-1} a_j^2 c_{n-1j})a_{n-1}^2 &= 0. \end{aligned} \tag{2.B.13}$$

20

The solution to this set lies on the conical surface defined in (2.B.12) and therefore has more than one nonzero elements. Solutions other than $\mathbf{a} = \mathbf{0}$ can be found by letting $t \geq 2$ of the $n-1$ variables $a_1, a_2, \ldots, a_{n-1}$ be nonzero and letting

$$K = \{k_i : a_{k_i} \neq 0, \ i = 1, 2, \ldots, t\} \tag{2.B.14}$$

be an index set. From (2.B.13) we obtain the following system of equations by dividing equation $k_i \in K$ by $a_{k_i}^2$

$$\begin{bmatrix} 0 & -c_{k_1,k_2} & -c_{k_1,k_3} & \cdots & -c_{k_1,k_t} \\ -c_{k_2,k_1} & 0 & -c_{k_2,k_3} & \cdots & -c_{k_2 k_t} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ -c_{k_t,k_1} & -c_{k_t,k_2} & \cdots & -c_{k_t,k_{t-1}} & 0 \end{bmatrix} \begin{bmatrix} a_{k_1}^2 \\ a_{k_2}^2 \\ \vdots \\ \vdots \\ a_{k_t}^2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_t \end{bmatrix}$$

$$\tag{2.B.15}$$

which is linear in $a_{k_1}^2, a_{k_2}^2, \ldots, a_{k_t}^2$. Rewriting the last equation as

$$\mathbf{C}\mathbf{y} = \mathbf{b} \tag{2.B.16}$$

where $\mathbf{y} = [a_{k_1}^2 \ a_{k_2}^2 \ \cdots \ a_{k_t}^2]$. One can show that $rank(\mathbf{C}) = rank([\mathbf{C}|\mathbf{b}]) = 2$ for any $t \geq 2$. Therefore, the solution to (2.B.16) is not unique. Among all possible solutions, we seek those that are feasible, i.e., lie inside the unit sphere. Applying elementary row operations to the augmented matrix $[\mathbf{C}|\mathbf{b}]$ we obtain the following 2 constraints

$$\begin{bmatrix} 0 & (x-1)^2(\lambda_{k_1} - \lambda_{k_2}) & (x-1)^2(\lambda_1 - \lambda_{k_3}) & \cdots & (x-1)^2(\lambda_{k_1} - \lambda_{k_t}) \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{k_1}^2 \\ a_{k_2}^2 \\ \vdots \\ a_{k_t}^2 \end{bmatrix}$$

$$= \begin{bmatrix} x^4(\lambda_{k_1} - \lambda) - 2x^3\lambda_{k_1} - x^2(\lambda_{k_1} + \lambda) \\ x^2 \end{bmatrix}. \tag{2.B.17}$$

These constraints must be satisfied by any solution to (2.B.11). It is apparent from the second constraint that $\sum_{j=1}^{t} a_{k_j}^2 = x^2$. But since we assumed that $x > 1$, and $\sum_{j=1}^{t} a_{k_j}^2 = x^2$, any solution to (2.B.11) other than the origin, $\mathbf{a} = \mathbf{0}$, is therefore infeasible.

We have thus found the only stationary point of $f(\mathbf{a}, \lambda)$ when $\mu = 0$. To see when $\bar{\mathbf{a}}^1 = \mathbf{0}$ is a maximum point for $f(\mathbf{a}, \lambda)$ we examine the second order sufficient conditions as follows:

$$\nabla^2 L(\bar{\mathbf{a}}^1) = \begin{bmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & b_{n-1} \end{bmatrix}. \tag{2.B.18}$$

If all the b_i's have the same sign, $\mathbf{a} = \mathbf{0}$ is either a maximum point if the sign is negative or a minimum point if the sign is positive. If there is a sign change $\mathbf{a} = \mathbf{0}$ is a saddle point. The point $\bar{\mathbf{a}}^1 = \mathbf{0}$ is a maximum when all the coefficients b_i, $i = 1, \ldots, n-1$ are negative. From the definition of b_i in (2.B.9) this occurs when $x = 1 + \sigma_v^2/\sigma_w^2 < (1 + \lambda/\lambda_1)/(1 - \lambda/\lambda_1)$. When $\bar{\mathbf{a}}^1 = \mathbf{0}$ corresponds to a maximum, it follows from (2.B.4) that $\mathbf{s}_n = a_n \mathbf{s}^\perp$. Since $\sum_i^n a_i^2 = 1$, $a_n = 1$ and therefore $\mathbf{s}_n \in V^\perp$. We have established that the potential signal to transmit at stage n is orthogonal to all previously transmitted signals. To determine an explicit form for $\mathbf{s}_n$ in this case, we substitute with $\mathbf{a} = \mathbf{0}$ into (2.B.5) and maximize with respect to λ. One can easily see that the maximum occurs when $\lambda = \lambda_n$, the the nth largest eigenvalue of $\bar{\mathbf{A}}^T \bar{\mathbf{A}}$. The signal $\mathbf{s}_n$ is therefore the eigenvector of $\bar{\mathbf{A}}^T \bar{\mathbf{A}}$ corresponding to λ_n.

2.B.B.2 Constrained Stationary Points: $\mu > 0$

This case requires the constraint $\mathbf{a}^T\mathbf{a} = 1$ to be active at the solution point. We can proceed as we did for $\mu = 0$ by determining the KKT points for this case. Alternatively, we could find the maximum point of $f(\mathbf{a}, \lambda)$ on the unit sphere $\|\mathbf{a}\| = 1$ by

$$\max_{\|\mathbf{a}\|=1} f(\mathbf{a}, \lambda) = \max_{\|\mathbf{a}\|=1} \frac{\frac{1}{\sigma_w^2}[x^2(\lambda + \sum_{k=1}^{n-1} \lambda_k) + \mathbf{a}^T\mathbf{Q}\mathbf{a}]}{x[x^2 - \mathbf{a}^T\mathbf{a}]} \qquad (2.B.19)$$

$$\leq \frac{\frac{1}{\sigma_w^2}[x^2(\lambda + \sum_{k=1}^{n-1} \lambda_k) + q_{11}]}{x[x^2 - 1]} \qquad (2.B.20)$$

where the last inequality followed from the fact that $q_{11} \geq q_{22} \geq \cdots \geq q_{n-1,n-1}$ and that $\mathbf{a}^T\mathbf{a} = 1$. Equality in (2.B.20) occurs when $\mathbf{a} = [1 \; 0 \; \cdots \; 0]$ at which point $f(\mathbf{a}, \lambda)$ attains its maximum. Therefore, for the case $\mu > 0$, we have one potential global maximum point.

2.B.B.3 Global Optimality Conditions

The global optimality conditions are determined by comparing the value of the function $f(\mathbf{a}, \lambda)$ at $\bar{\mathbf{a}}^1 = \mathbf{0}$ and $\bar{\mathbf{a}}^2 = [1 \; 0 \; \cdots \; 0]$. From (2.B.5), one can see that

$$f(\bar{\mathbf{a}}^1, \lambda) = \frac{1}{\sigma_w^2} \frac{\lambda + \sum_{k=1}^{n-1} \lambda_k}{x}, \qquad (2.B.21)$$

$$f(\bar{\mathbf{a}}^2, \lambda) = \frac{1}{\sigma_w^2} \frac{x(\lambda_1 + \sum_{k=1}^{n-1} \lambda_k) + \lambda_1 - \sum_{k=1}^{n-1} \lambda_k}{x(x - 1)}. \qquad (2.B.22)$$

One can show that when $x < \frac{1 + \lambda/\lambda_1}{1 - \lambda/\lambda_1}$, $f(\bar{\mathbf{a}}^1, \lambda) > f(\bar{\mathbf{a}}^2, \lambda)$. When this is the case, the signal to transmit at the nth stage $\mathbf{s}_n^*$ is the the eigenvector of $\bar{\mathbf{A}}^T\bar{\mathbf{A}}$ corresponding to λ_n. Otherwise, $\mathbf{s}_n^*$ should be a repetition of the first transmitted signal $\mathbf{s}_1^*$. If we decide to retransmit the first signal, we can obtain the new $KLIN$ from (2.B.5) by the substitution $\bar{\mathbf{a}} = [1 \; 0 \; \cdots \; 0]$ as follows:

$$KLIN^{(n)} = \frac{\lambda_1}{1 + \frac{\sigma_v^2/2}{\sigma_w^2}} + \frac{\sum_i^{n-2} \lambda_i}{1 + \sigma_v^2/\sigma_w^2}. \qquad (2.B.23)$$

Note that the retransmission of $\mathbf{s}_1^*$ increased the $KLIN$ by reducing the effect of noise in our measurement from the signal $\mathbf{s}_1^*$.

We have thus shown that the nth transmitted signal is determined by the ratio σ_v^2/σ_w^2. If $\sigma_v^2/\sigma_w^2 < 2(\frac{\lambda_n}{\lambda_1})/(1 - \frac{\lambda_n}{\lambda_1})$, then more information about the target type can be obtained if the nth signal is orthogonal to the previous $n - 1$ signals. If $\sigma_v^2/\sigma_w^2 > 2(\frac{\lambda_n}{\lambda_1})/(1 - \frac{\lambda_n}{\lambda_1})$, then it is better to retransmit the first signal.

2.B.C General Case

We started our search for the optimal signal by assuming that all previously transmitted signals were orthogonal. We now relax this condition and assume that retransmissions have occurred and that at the $n - 1$st stage the signals $\mathbf{s}_1^*, \mathbf{s}_2^*, \ldots, \mathbf{s}_j^*$ were repeated $n_1, n_2, \ldots, n_j$ times, respectively. Since $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_j$, we must have

$$n_1 \geq n_2 \geq \cdots \geq n_j. \qquad (2.B.24)$$

Let the signal to transmit at stage n be given by

$$\mathbf{s}_n = a_1\mathbf{s}_1^* + a_2\mathbf{s}_2^* + \cdots + a_j\mathbf{s}_j^* + a_{j+1}\mathbf{s}^\perp \qquad (2.B.25)$$

where $s^\perp$ is orthogonal to $\mathbf{W} = span\{s_1^*, s_2^*, \ldots, s_j^*\}$. Following steps similar to the ones that led to (2.B.3) and (2.B.5), one can show that

$$\vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \vec{s} = \frac{\frac{1}{\sigma_w^2}[(1 + \frac{\sigma_v^2}{\sigma_w^2})(\lambda' + \sum_{i=1}^{n-1} \lambda_i') + \mathbf{a}^T \mathbf{Q}' \mathbf{a}]}{[(1 + \frac{\sigma_v^2}{\sigma_w^2}) - \mathbf{a}^T \mathbf{D} \mathbf{a}]} \tag{2.B.26}$$

where the matrix $\mathbf{Q}' = [q_{ii}']$ is diagonal with entries

$$q_{ii} = x_i^2(\lambda_i' - \lambda^{(i)}) - 2x_i\lambda_i' - \sum_{j \neq i} \lambda_j' \tag{2.B.27}$$

and

$$x_i = (1 + \frac{\sigma_v^2/n_i}{\sigma_w^2}), \qquad \lambda_i' = \frac{\lambda_i}{x_i}, \qquad \lambda^{(i)} = \frac{\lambda}{x_i}, \qquad \lambda' = \frac{\lambda}{(1 + \frac{\sigma_v^2}{\sigma_w^2})}. \tag{2.B.28}$$

The matrix $\mathbf{D}$ is also diagonal with entries $d_{ii} = \frac{1}{x_i}$. The maximization of (2.B.26) proceeds along the same lines as the maximization of (2.B.5).

2.B.C.1   Unconstrained Stationary Points: $\mu = 0$

By setting up the Lagrangian function as in (2.B.7), we can show that the only feasible solution with $\mu = 0$ is $\bar{\mathbf{a}}^1 = 0$. This point is a maximum of (2.B.26) when

$$b_i = (1 + \frac{\sigma_v^2}{\sigma_w^2})(x_1^2(\lambda_i' - \lambda^{(1)}) - 2x_1\lambda_i' + (\lambda_i' + \lambda')) < 0, \qquad i = 1, 2, \ldots, n-1. \tag{2.B.29}$$

The point $\bar{\mathbf{a}}^1$ corresponds to a selection of a signal from the subspace $\mathbf{W}^\perp$. By substituting with $\bar{\mathbf{a}}^1$ in (2.B.26) and maximizing with respect to λ the signal to transmit would be s_{j+1}, the eigenvector of $\bar{\mathbf{A}}^T \bar{\mathbf{A}}$ corresponding to the $(j+1)$st eigenvalue. The discrimination defined in 2.4.40 gain for this case is given by

$$\Delta KLIN^\perp = \frac{\lambda_{j+1}}{1 + \frac{\sigma_v^2}{\sigma_w^2}} \tag{2.B.30}$$

where the superscript $\perp$ denotes that the signal s_{j+1} is orthogonal to $\mathbf{W}$.

2.B.C.2   Constrained Stationary Points: $\mu > 0$

The case when $\mu > 0$ can also be examined by maximizing $f(\mathbf{a}, \lambda)$ in (2.B.26) on the unit sphere $\|\mathbf{a}\| = 1$ as before. Since the solution lies on the unit sphere, it must come from $\mathbf{W}$. Equation (2.B.26) can also be written as

$$\vec{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \vec{s} = \sum_{i=1}^{j} \frac{\lambda_i}{1 + \frac{\sigma_v^2/n_i}{\sigma_w^2}} + \sum_{i=1}^{j} \frac{\frac{\sigma_v^2/n_i}{\sigma_w^2} \lambda_i'}{[(1 + \frac{\sigma_v^2}{\sigma_w^2}) - \mathbf{a}^T \mathbf{D}\, \mathbf{a}]} \frac{a_i^2}{x_i}. \tag{2.B.31}$$

The second part of the right handside of (2.B.31) is the discrimination gain due to some signal in $\mathbf{W}$ and can be written as

$$\Delta KLIN^\| = \frac{\mathbf{a}^T \mathbf{E}\, \mathbf{a}}{\mathbf{a}^T (x\mathbf{I} - \mathbf{D})\, \mathbf{a}} \tag{2.B.32}$$

where $\mathbf{E}$ is diagonal with entries $e_{ii} = \frac{\lambda_i'}{x_i} (\frac{\sigma_v^2/n_i}{\sigma_w^2})$ and the superscript $\|$ denotes that the discrimination gain is due to a signal in $\mathbf{W}$. The maximization of (2.B.26) on the unit sphere $\|a\| = 1$ is then equivalent to maximizing (2.B.32). Maximizing (2.B.32) on the unit sphere is a generalized eigenproblem where the solution is the unit norm eigenvector corresponding to the maximum eigenvalue of the matrix

$$(x\mathbf{I} - \mathbf{D})^{-1}\mathbf{E}. \tag{2.B.33}$$

23

Since this matrix is diagonal, the eigenvectors are the standard basis vectors. The maximizing eigenvector is the vector $\bar{a}^2 = [0 \cdots 0\ 1\ 0 \cdots 0]$ where the nonzero entry corresponds to the largest entry in the matrix of (2.B.33) and that is given by

$$\max_{1 \le i \le j} \frac{\lambda_i'(\frac{\sigma_v^2/n_i}{\sigma_w^2})}{(n_i + 1) + \frac{\sigma_v^2}{\sigma_w^2}}. \tag{2.B.34}$$

2.B.C.3 Global Optimality Conditions

Global optimality conditions are determined by examining the discrimination gains $\Delta KLIN^\perp$ and $\Delta KLIN^\parallel$ in (2.B.32) with $\bar{a}^2$. If

$$\frac{\lambda_{j+1}}{1 + \frac{\sigma_v^2}{\sigma_w^2}} > \max_{1 \le i \le j} \frac{\lambda_i'(\frac{\sigma_v^2/n_i}{\sigma_w^2})}{(n_i + 1) + \frac{\sigma_v^2}{\sigma_w^2}} \tag{2.B.35}$$

then transmit the $(j+1)$st eigenvector of $\bar{A}^T \bar{A}$ otherwise retransmit the kth eigenvector where

$$k = \arg \max_{1 \le i \le j} \frac{\lambda_i'(\frac{\sigma_v^2/n_i}{\sigma_w^2})}{(n_i + 1) + \frac{\sigma_v^2}{\sigma_w^2}}. \tag{2.B.36}$$

Chapter 3

Image Coding for Query by Pictorial Content

3.1 Introduction

The need for sophisticated data management techniques continues to grow with the proliferation of very large image databases, e.g., online digital libraries, digital art collections, biomedical image libraries, merchandise catalogs, satellite imagery, fingerprint and mug-shot archives, etc. Due to their size, most need to be stored in a compressed form to conserve storage space and delivery bandwidth. At the same time, some of the most important uses of such a collection are retrieving, manipulating, and browsing the stored images in terms of *image content*, e.g., colors, shapes, textures, etc. A schematic of a typical content-based query system is shown in Fig. 3.1. A user with an information need presents an example image object, texture swatch, or sketch and requests similar images from the image collection.

To be effective, the image management system should provide

- efficient storage of the image collection,

- fast content-based searching of the images, and

- user-friendly browsing of the database and retrieval results.

These allow very large image collections to be stored while simultaneously helping users (even those unfamiliar with the database) retrieve relevant images based on objects, shapes, textures, colors, etc.

The traditional image management approach is to handle compression and retrieval *separately*. The reason for this is that traditional compression techniques are oriented towards storage and transmission efficiency. They do not address the issue of retrieving files via content-based searches. On the other hand, the indexing methods used for retrieval concentrate on providing a structure to match queries. Usually, an index is a separate file or file header which provides marks or pointers to instances of terms in the database. No compression of the original data is obtained. In fact, compression and indexing are usually at odds. Without an index, compression makes content-based retrieval more difficult since the data usually must be decompressed before matching can be performed. With an index, additional storage space must be allocated to maintain the index. This diminishes the benefit of compressing the original images. A diagram depicting this problematic situation is shown in Fig. 3.2(a) and (b). Such a data management system is inefficient and often redundant, since the information contained in the content-based index consists mainly of data taken or derived from the file collection.

We propose a new data representations which combines compact coding with support for content-based retrieval. In particular, we present a new coding algorithm to accomplish this task. In Fig. 3.2(c), we show the efficient storage requirements of this new data management technique. The approach implicitly indexes a file collection by building query support directly into the compressed files. Furthermore, the algorithm has several strong features which are particularly useful for an advanced data management system:
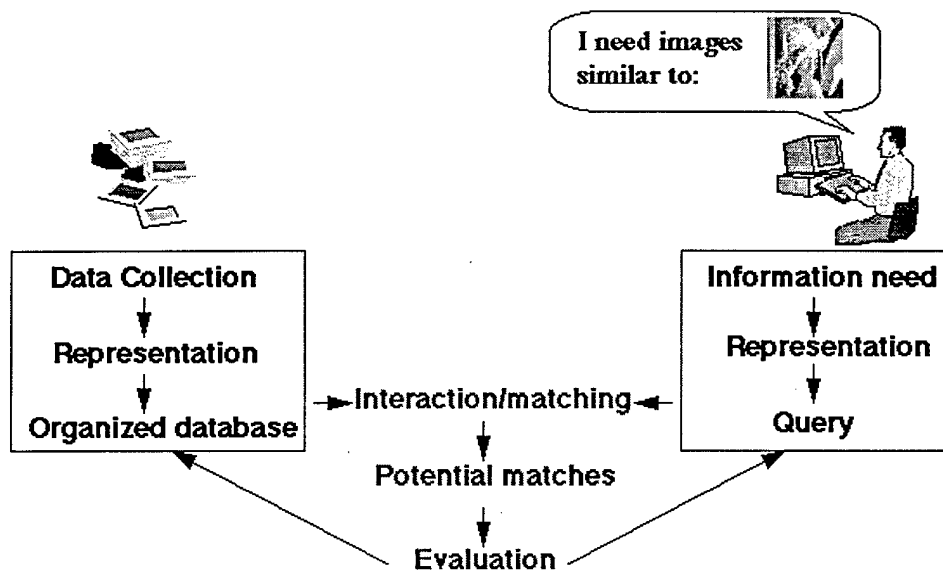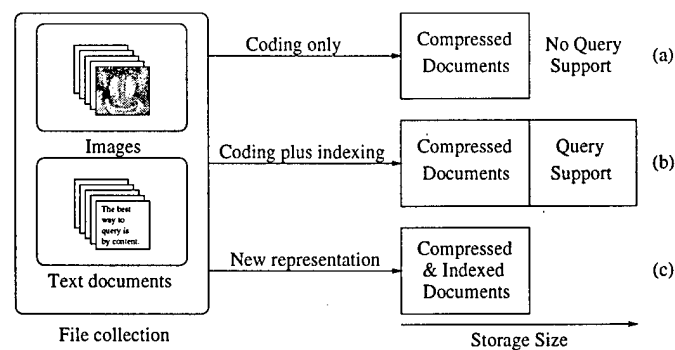
Figure 3.1: Content-based retrieval scenario.

Figure 3.2: Storage requirements for a file collection in terms of (a) compressed files, (b) compressed files and separate index, and (c) our new coding technique.

- *flexibility:* can minimize a weighted sum of the expected file size in bits (compressed file size) and the expected number of bits that need to be read to answer a query (query response time). In a system where storage space is scarce, emphasis may be placed on compressing the data. In applications where query speed is important, the system can sacrifice some compression to concentrate on retrieval.

- *universality:* applicable to any data in which objects or primitives can be defined. We will describe the system for document and images, although audio and video are also possible. This property is very useful for the wide range of multimedia data types.

- *compressed data modification:* allows manipulation and modification of compressed data. This is much more efficient as the data, when compressed, requires fewer operations and less memory.

- *retrieval based on bit patterns:* no expensive computations (e.g., Euclidean distance) are required during retrieval. Retrieval of image files relies solely on bit pattern comparisons.

- *progressive refinement retrieval:* successively reduces the number of searched files as more bits are read. Based on multiresolution techniques, each file is coded in terms of a coarse-to-fine discriminant structure. At the coarsest scale, only a few bits are read and a large number of files not matching the query criteria are quickly rejected. Searching the remaining documents proceeds in terms of the next discriminant scale. This offers a significant speed improvement by quickly rejecting files as potential retrievals and concentrating search resources on fewer files.

- *high quality data browsing at low bit rates:* based on embedded prototypes, our approach supports high quality browsing even when a minimum number of query bits are read. Rather than retrieving and displaying images in blurred form (which is typical for progressive transmission systems), image objects are represented in full detail at each stage in the retrieval. Specifically, after a given number of bits are read, the images most similar to a query in the collection are returned and displayed with full detail prototype objects representing objects within the image.

The coding algorithm we propose may be directly applied to compress existing separate indexes (see Sec. 3.3) without compressing the data files. However, the approach offers considerable benefit when used to simultaneously compress and index a file collection. In what follows, we present our new integrated data representation, address the advanced database issues described above, and include retrieval and browsing examples.

3.2 Overview of new data representation

Our new representation is achieved by first extracting all the important information, i.e., query terms, from the document we wish to compress and index. For textual documents, query terms are words and phrases. These are the textual objects of interest. In a similar fashion, we define objects of interest for image retrieval in terms of image objects and regions, e.g., geometrical shapes, faces, trees, clouds, etc. In this way, the approach is consistent with the new generation of object-based coders (like MPEG-4) and other similarity retrieval systems (see Sec. 3.3). A multiresolution representation is then computed for each query extracted from the document. Our new coding algorithm then assigns *codewords* to the multiresolution representation of each query term. The algorithm also specifies a *relative position* for each query term codeword in the coded file. Codewords and position are obtained using the probability that the object occurs in a document and the probability the object will be queried. The text or image file is coded into three sections: 1) a file header consisting of concatenated query term codewords, 2) a set of indices denoting the locations of these terms in the file, and 3) the remaining non-query terms in the file. This is shown in Fig. 3.3. Using the ordering specified by the algorithm, each file header is constructed by concatenating the codewords of a multiresolution representation of each query term which appear in that file.

The compressed file header contains all of the information needed to answer a content-based query. When a query is posed, the compressed file header of each document or image is searched sequentially for the query term codewords (i.e., bit patterns) instead of the original uncompressed file or a separate index. Since the relative order of the terms is known *a priori*, a search is terminated early once we read a codeword that corresponds to a query term that should appear after that corresponding to the actual query.
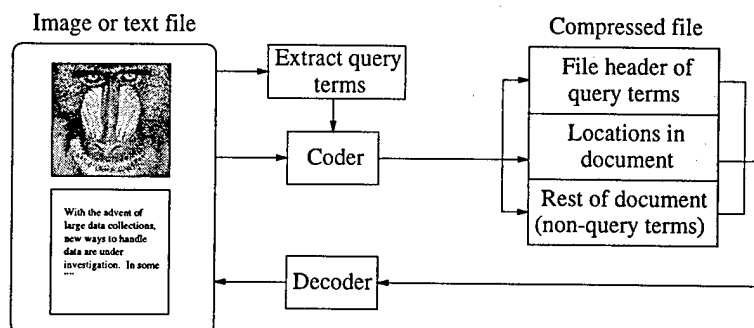
Figure 3.3: Diagram of new coding technique showing three sections of compressed file.

To reconstruct a file from its coded version, the non-query data in the file are decompressed. Each term in the file header is then decoded and inserted into the main part of the document or image according to the coded term locations. Thus, there is no ambiguity during the decoding process.

3.3 Previous work

Traditional image retrieval systems are based on keywords and captions rather than content [28]. User supplied keywords are stored along with the compressed images. A query, "show me all images that have red cars in them," retrieves images by scanning the textual information appended to each image in the database. Although useful for some purposes, it is difficult to capture visual or image properties with textual descriptors. As a result, methods that address image retrieval based on image content are needed [29]. Recently, content-based image retrieval systems have received a great deal of attention in the literature, e.g., the special issue on content-based image retrieval systems of *IEEE Computer*, September 1995. We review here some of the pioneering techniques that have been proposed.

IBM has developed the Query by Image Content (QBIC) system to explore content-based image retrieval methods [30, 31]. QBIC is commercially available in IBM's Ultimedia Manager software package. The QBIC system is based on feature vectors, which are numerical attributes constructed from image properties (e.g., colors, shapes, textures, transform coefficients). Feature vectors are used to describe an image in terms of content. During database population, a user manually or semi-automatically identifies regions of interest (i.e., objects) in the images. Feature vectors are computed for *each image* and *each image object identified during database population* and stored as side information with the image data.

A query in QBIC may be posed using example images, user-constructed sketches, and color/texture patterns. In a query, features from the database are compared to corresponding features from the query specification to determine which images are a good match. Image similarity is defined in terms of a distance measure between corresponding feature vectors. Note that QBIC, like most content-based query approaches, uses *degree of similarity* rather than exact match to define similarity between the query term and the images in the database. This is unlike most text database systems which process queries based on exact match. After computing the degree of similarity for each image in the database, the N most similar images to the query are displayed to the user.

The Massachusetts Institute of Technology has developed a set of interactive tools for browsing and searching images called "Photobook" [32]. Photobook supports content-based image retrieval by using semantics preserving image compression. Photobook allows images to be searched on appearance, 2D shape, and texture.

Photobook's semantics preserving image compression is a collection of transforms which are used to reduce an image or image object to a small set of perceptually-significant coefficients. Photobook uses a variety of image representations for different types of image objects and regions. For objects like faces and eyes, Photobook uses a Karhunen-Loeve transform (KLT) representation. Objects which are more texture-based are represented using a Wold decomposition. A Wold decomposition represents a texture in terms of three orthogonal components: harmonics, directionality, and noise. The transform coefficients (either Wold or KLT) of each object in an image are stored and used for similarity comparison. During retrieval, the

transform coefficients of the query object are obtained and compared with images stored in the database. Their approach is similar to ours in that it builds query support directly into the coded image representation. Specifically, the transforms (like the discrete cosine transform used in JPEG [33]) are complete and may be used to represent the image object.

Another technique uses a signature based scheme for image querying [34]. The images in the database are stored using a standard compression algorithm (e.g., JPEG, GIF, wavelet coder). Separate image *signatures* are stored for retrieval. A signature is created for each image in terms of wavelet coefficients. Specifically, an image signature is constructed by first applying a Haar wavelet decomposition to each YIQ color component of the image. Then, for each of the three color channels, the m largest magnitude coefficients are kept. Each of these coefficients is quantized to one of two levels: $+1$ representing large positive coefficients; or -1 representing large negative coefficients. The overall signature for each image consists of the image's overall average color and the indices and signs of its m largest magnitude wavelet coefficients. The indices for all of the database images are organized into a set of data structures (inverted file indexes) for searching. For each query image, the same wavelet transform is applied with only the average color and m largest wavelet coefficients kept. The retained information is used for similarity comparisons.

In [35], a technique for browsing large-scale aerial photographs is presented. The system works by building a texture thesaurus model for fast search and indexing. The texture features are computed by filtering the image with a bank of Gabor filters. This is followed by texture flow computation to segment each large airphoto into homogeneous regions. The overall system is used for searching over a large collection of airphotos for geographic features such as housing developments, parking lots, highways, and airports. Content-based retrieval systems are also proposed in [36] and [37]. In [37], an image object is represented by a set of structural components which represent a point in multidimensional space. A retrieval scheme which represents images using pseudo two-dimensional hidden Markov models is proposed in [38].

In general, all of these content-based retrieval schemes and several others provide significant advantages over keyword retrieval. Unfortunately, any techniques based on *storing side information* for retrieval are inefficient. In the QBIC system, for example, the numerous feature vectors appended to each image may require as much storage space as the images themselves!

The current systems also suffer when it comes to query response speed. First, storing side information not only increases storage requirements but also increases response time. In particular, the appended search information must be read and processed during the retrieval. As the side information increases the amount of data increases. Furthermore, since image similarity must be computed during the retrieval, care must be taken to maintain high speed retrieval when computing similarity between high dimensional feature vectors in very large image collections. A typical distance metric between two length N vectors requires $2N$ additions and N multiplications. For example, retrieval by color using a length $N = 256$ histogram for a small database of 10,000 images and assuming 5 objects per image requires 10^7 additions and multiplications *per query*! A great deal of research effort focuses on this dilemma.

In addition to similarity indexing, methods to manipulate and retrieve compressed data are also under investigation. These methods, like ours, access the data in compressed form. Techniques to extract and manipulate compressed image data are presented in [39]. In [40], the authors propose finding regions of interest (i.e., query terms) within JPEG and MPEG coded data using only the length of the compressed DCT coefficients. These methods and others provide a significant reduction in computational complexity compared to processing uncompressed forms. However, the bit streams from JPEG and other coding algorithms are difficult to search and the approaches are largely heuristic. The bit stream produced by our coding algorithm is designed to take searching into account.

3.4 Combining compression and retrieval

The key idea of our representation is the following: given any set of terms which have a probability of occurring in a file and a probability of being queried, we can code the terms to minimize a weighted sum of the expected compressed file size and expected query response time.

We begin by discussing our approach of integrating compression and indexing in a text document environment. The coding algorithm works similar to the example shown in Fig. 3.4. In the example, three query terms have been extracted from the text document and placed in the file header. Locations of the query
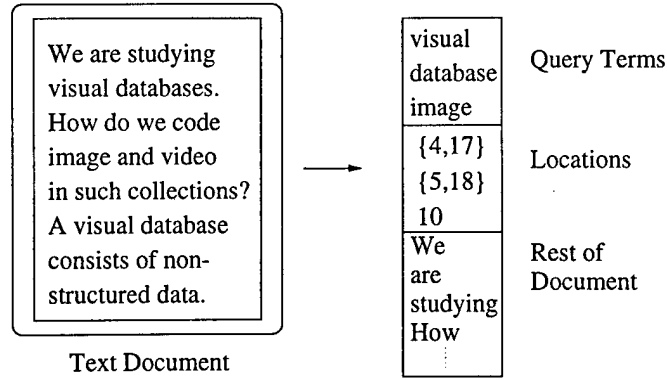
Figure 3.4: Example of a text document, query terms (words) and locations of the query terms.

terms and the remaining portion of the document are also stored. This information must be retained so that the original document can be reconstructed. When a query is posed, the header is searched sequentially for the query term. For example, the query "get all documents with 'visual' in them" would result in a match for the example document.

In our example, the important terms in the document have been grouped but not compressed or ordered for fast retrieval. Compression can be achieved by exploiting term probabilities. In particular, we can use entropy coding to represent common terms with short codewords and rare terms with longer codewords. To achieve fast retrieval, the query terms in the header must be properly ordered. Clearly, if we plan to search the header sequentially for query terms (i.e., bit patterns), we should put terms which are frequently queried near the beginning of the header. That way we can often determine whether a document contains a given term quickly. Both of these issues are addressed by our approach.

Our algorithm assigns a codeword and relative ordering to each query term in the file to minimize a weighted sum of the expected compressed file size and the expected query response time. The weighted sum we wish to minimize can be expressed as

$$C = E[\text{Search Length}] + \lambda E[\text{File Length}], \qquad (3.1)$$

where $E[\cdot]$ is the expected value. The user is free to choose the weight $\lambda \geq 0$ to control the tradeoff between compression and query response. For $\lambda < 1$, the emphasis of the data representation is better query response time. For $\lambda > 1$, emphasis is placed on minimizing expected file length. This flexibility allows the algorithm to adapt to the needs of a given database environment. In Appendix 3.A and in [41], we define the cost function C in terms of the probability and query probability distributions of the query terms in the database. Furthermore, we show how it may be *minimized* by assigning codewords and ordering to the query terms.

In addition to reducing storage overhead and increasing retrieval speed, our coding method has other distinct advantages. First, it has built in support for a nearness constraint between two or more query terms as the location of each query term is saved in the coded file. For example, two terms in a query may be required to occur adjacent to or within N words of each other. Second, the new coding technique directly supports modification or manipulation of the coded data. For example, we can perform an efficient "find and replace" operation on the compressed data.

3.5 Combining compression and retrieval for images

While described in terms of text documents, the integrated coding approach is readily applicable to image management. Image retrieval is based on properties of image regions and objects. In this section, we describe how image objects are defined and coded using our new coding algorithm. A description of our image coding algorithm may be found in Appendix 3.B and in [42].
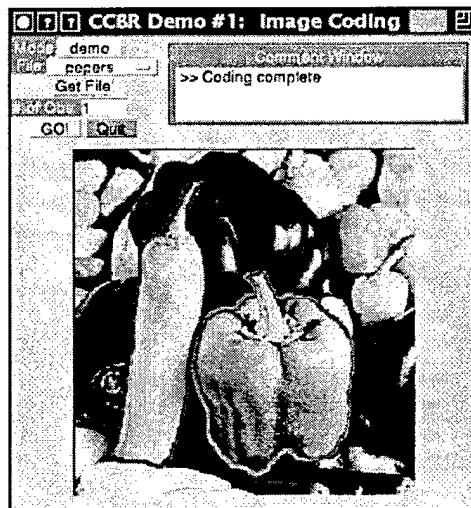
Figure 3.5: Defining two objects (peppers) in an image.

3.5.A   Defining objects and processing

In a text database, query terms (words) are relatively easy to define and extract. In an image database, we define query terms as objects and regions within the image. Image regions and objects may denote any entities of interest in the given database application. Examples include shape-based objects like faces, eyes, cars, buildings, windows, etc. Other examples include items more texture- or color-based than shape-based, e.g., regions of grass, sand, clouds, trees, etc. To extract the query terms, image regions and objects must be identified. Ideally, this process would be done automatically using a segmentation algorithm. This is currently an unresolved issue. In our implementation, objects are manually defined using a graphical user interface and a mouse. The result of defining two objects in an image is shown in Fig. 3.5. Defining objects is further described in Appendix 3.B.

The next step consists of finding an object representation that is useful for matching. We use a wavelet transform for this purpose since it is capable of compactly representing each object at different scales and resolutions [43]. The output of the wavelet transform is a collection of subbands ranging from low to high resolution which completely describe the original object. The subband decomposition for the bell pepper shown in Fig. 3.5 is shown in Fig. 3.6.

An inherent property of the representation is that coarser subbands contain fewer transform coefficients than finer subbands. Note that we could use other object representations (e.g., KLT or Wold transforms used by MIT's Photobook). As shown in Sec. 3.6, the results obtained using the wavelet representation are promising.

After the objects are extracted from the original image, the remaining non-object image regions are coded using a simple block coding algorithm like JPEG [33]. The coded background is stored in the last part of the encoded file (c.f. Fig. 3.3). In addition, the position and size of each object is stored for decoding purposes.

3.5.B   New image representation

Once objects are defined and transformed as above, our new coding technique is applied. The coder constructs the file header in terms of the wavelet transform subbands of the segmented objects to minimize the size-search cost function (Eq. 3.1). To obtain the probability and query probability estimates used by our coding algorithm, we use vector quantization (VQ) [44] to map the subbands to finite dictionaries. Using the techniques developed in Section 3.4, the probabilities associated with each term in the dictionary determine the proper location and codeword of each object subband in the compressed file header. To increase retrieval speed, the file header is further divided by resolution. Specifically, the image file header is ordered in a coarse-to-fine manner.
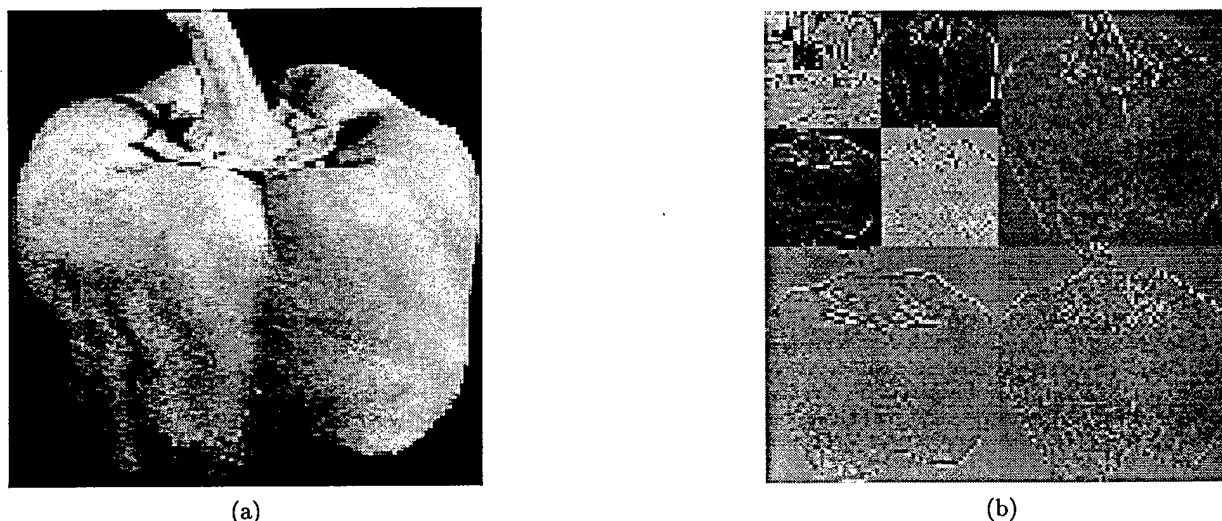
31

(a)	(b)

Figure 3.6: Subbands of bell pepper object, (a) original object, and (b) wavelet transform subbands. The upper left (lower right) subband corresponds to lowest (highest) frequency data.

3.5.C Search and retrieval of images

To initiate a search, user-defined queries, in the form of sample image objects, are applied as an input to the wavelet transform. The resulting query term subbands are mapped to the finite VQ dictionaries to obtain the appropriate codewords (i.e., bit patterns) for the search. Note that, once the query term subbands are mapped to VQ dictionaries, no computations are required during the retrieval. A search begins by searching the coarsest subband codewords of each image in the database for the query term codeword bits. If the bit pattern is matched, we have found an image with a similar object at that scale. Since the relative order of the codewords is known *a priori*, a search is terminated early once we read a codeword that corresponds to a query term that should appear after that corresponding to the actual query. Files which do not match the query criteria are rejected. Searching proceeds on remaining documents in terms of the next coarsest subband. In this way, our new coding technique implements a progressive refinement retrieval by successively reducing the number of searched files as more bits are read. This approach significantly improves retrieval speed by quickly rejecting files as potential retrievals.

Note that manipulation and modification of image objects regions are supported as the objects are encoded in the file header. This is difficult to do with a management system based on a separate index without first decoding the data. Usually an index is just a copy of the data, not the image data itself. We can also completely decode an image object while leaving the rest of the image compressed. Also, the similarity measure may be location dependent or independent. Since the location of each object in the image is stored, the query specification may include an image location where the object should appear.

3.5.D Embedded prototypes

Our image coding algorithm supports two special kinds of progressive image transmissions for browsing. First, the image headers may be progressively decoded (coarse-to-fine) and placed in an image according to their locations. In this way, browsing of objects in an image is supported. Conventional progressive image transmission decodes a coarse-to-fine version of the entire image. Ours decodes a coarse-to-fine version of the objects in an image. For the same number of decoded bits, our decoder presents a much sharper version of important image objects (i.e., the query terms).

The second, more powerful, browsing technique represents image objects with high quality image object *prototypes*. Rather than displaying objects in a blurred form, this technique replaces image objects with full detail object prototypes, i.e., representatives, at each stage during the retrieval process. The prototype which represents an object during the retrieval process depends on the number of bits read to that point in the retrieval. The displayed objects are correctly placed and scaled within the retrieved image, but they
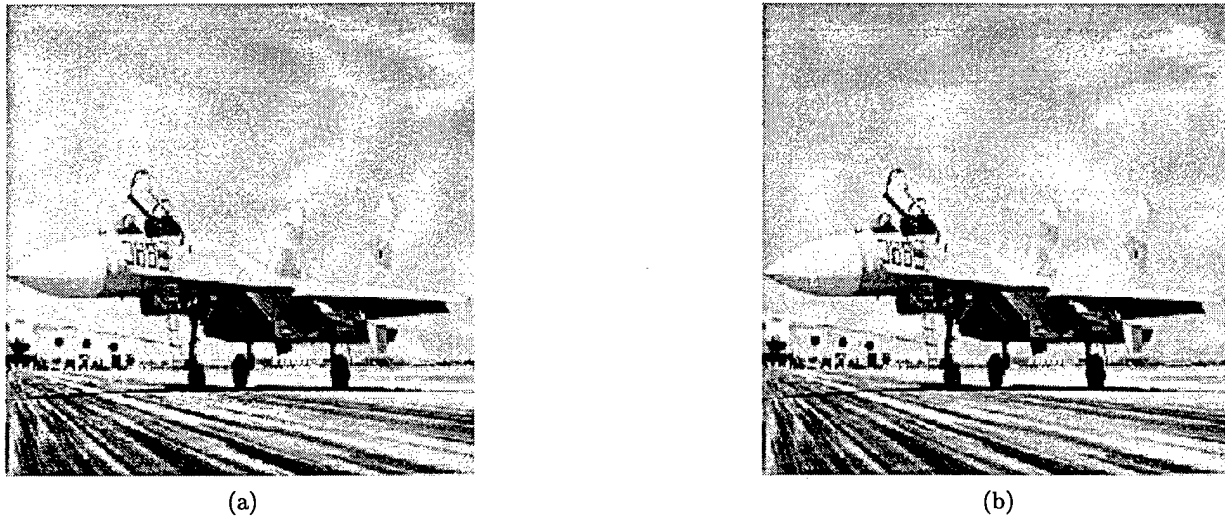
32

<div align="center">(a)</div> <div align="center">(b)</div>

Figure 3.7: Grayscale image of plane, (a) original image, and (b) image coded with new representation.

may be only an approximate (i.e., "low resolution" in terms of information) version of the actual objects. Note that the displayed prototype objects are high quality, i.e., real objects, not blurred objects. Multiscale prototypes are constructed from the previously described embedded VQ books shown in Fig. 3.16 and are defined in Appendix 3.B.

3.6 Examples of retrieval and browsing

To illustrate some of our techniques, we encoded a collection of 100 grayscale images of size 256×256. The collection includes a wide variety of images, e.g., faces, boats, fruit, planes, aerials, etc. Each image contains 4 to 8 objects of varying sizes. The number of objects in the database is about 500.

The compression performance of the algorithm indicates a slight increase (about 5-10%) in the size of the compressed files versus simply coding the files with JPEG. This seems very reasonable compared to storing a separate index. Image fidelity between our technique and the JPEG standard (Quality factor $Q = 75\%$) was chosen to be approximately the same. An example is shown in Fig. 3.7. The original image is shown in Fig. 3.7(a). After defining 5 objects in the image, the same image coded and stored using our new image representation is shown in Fig. 3.7(b). The images appear identical. Coding the original image using standard JPEG coding at quality factor $Q = 75\%$ results in a stored image size of 10158 bytes. The image stored using our technique requires 10988 bytes. The overhead in this case is 8.17%.

A second example is shown in Fig. 3.8(a) and (b). The original image is shown on the left. The coded file with 7 objects is shown on the right. Storing the image using JPEG coding requires 20076 bytes. The image stored with coding for retrieval required 21972 bytes. This amounts to a storage increase of 9.44%.

To illustrate the retrieval potential of this representation, we performed several queries. At retrieval time no similarity function (like Euclidean distance) needs to be computed. We are simply looking for a codeword (bit pattern) in the coded file. Since we know the relative position of the codewords *a priori*, a look-up table is consulted after each codeword is read to determine if the query codeword does not exist in the header. This allows searches to be terminated early. In Fig. 3.9 we show an example of a query object and two retrieved images ordered by similarity. In the query image, the user outlined the woman's face as the query object. As noted earlier, a wavelet representation of the query object is then computed and mapped to the finite VQ dictionaries to obtain codewords for each subband. Similarity is based on the number of subbands mapped to the same VQ terms as the query object. The objects outlined in the retrieved images were defined as query objects when they were coded and stored in the database. In this example, the query image exists in the database. The second retrieved image (next best match) is also the face of a woman.

In Fig. 3.10 we show a second example of a query object and two retrieved images ordered by similarity. The query term in this case is an airplane. The objects outlined in the retrieved images were defined as query
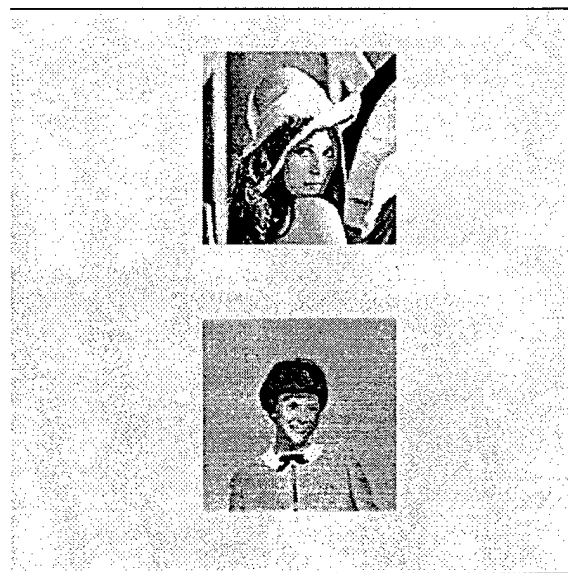
<div align="center">33</div>
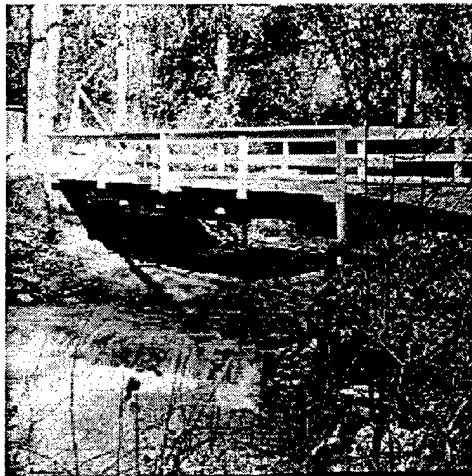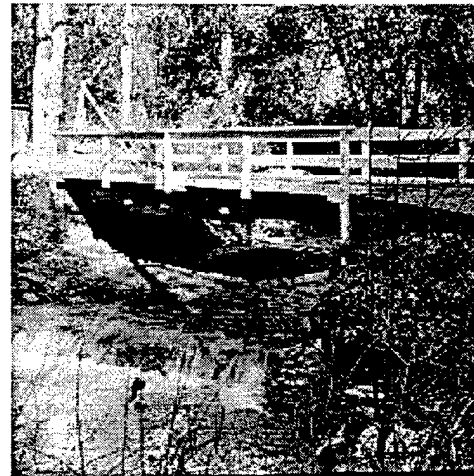
(a) (b)

Figure 3.8: Grayscale image of bridge, (a) original image, and (b) image coded with new representation.

(a) (b)

Figure 3.9: Retrieval example, (a) image with query object (face), and (b) two best retrieved objects.
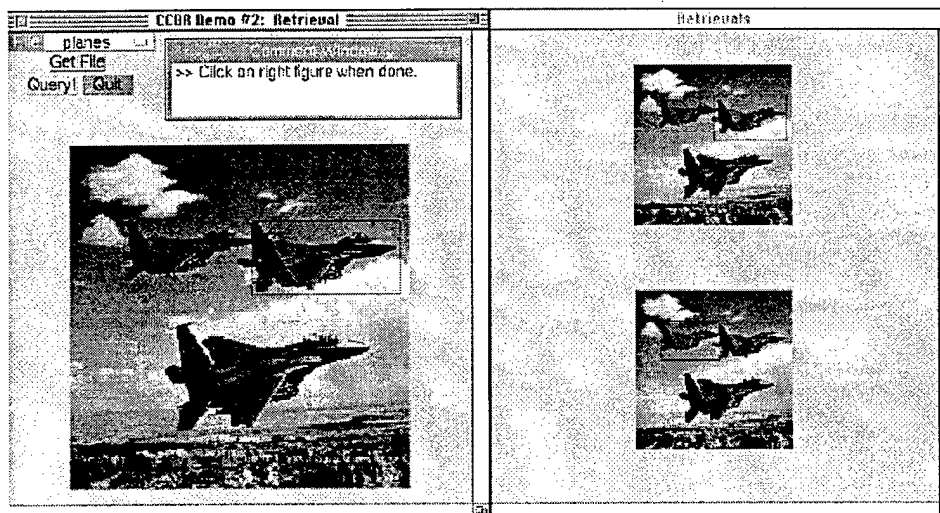
Figure 3.10: Original image and retrieved images.

objects when they were coded and stored in the database. Again, the query image exists in the database. Note, however, that when the image was originally coded and stored in the database, the plane object was defined somewhat differently than the user-defined query object. The similarity measure in this case is robust to some translation. Note too that the second retrieved image is the same image. However, in this case the plane object is a scaled and slightly rotated version of the original object.

In Fig. 3.11 we show a third example of a query object and two retrieved images ordered by similarity. Note that the retrieved image has a contour very similar to the query object.

We illustrate our browsing technique based on embedded prototypes on the plane retrieval in Fig. 3.10. Consider the images which could be displayed during the early stages of the retrieval. After reading the coarsest subband, the retrieved image could be displayed with the actual low resolution (i.e.,"blurred") plane object data as shown in Fig. 3.12(a). Another option is to use our embedded prototype approach. Using the VQ dictionaries, the prototype (i.e., representative) object is retrieved from the VQ codebook corresponding to the lowest resolution query codeword. Using the prototype, we replace the blurry data with the plane object's multiresolution prototype as shown in Fig. 3.12(b). The same number of bits are read during the retrieval for both images, yet the prototype image is easier to comprehend.

Appendix 3.A: Size-Search Cost Function

We define the cost function C in terms of the probability and query probability distributions of the query terms in the database. By assigning codewords and ordering to the query terms, we show how the cost function can be minimized.

Let us assume that we have one or more text files from which a set of n query terms $\{t_1, t_2, \ldots, t_n\}$ has been extracted. For each query term t_i, let p_{t_i} be the probability that the term exists in a file, and let the query probability q_{t_i} be the nonzero probability that term t_i occurs in a query.

The expected search length and expected file length depend on the length of the codewords assigned to each query term and the order with which the query terms appear in the coded file. For notational convenience, we introduce an *ordered* set of terms $\alpha_1, \alpha_2, \ldots, \alpha_n$. The subscripts on these symbols are used to denote the relative ordering of the terms when they occur in the file header, e.g., α_1 comes before α_2 when they both exist in a header. An ordering $\alpha_j = \Phi(t_i)$ associates each t_i with an α_j. Furthermore, let L_{α_i} be the codeword length of term α_i. Then the cost function we wish to minimize can rewritten as

$$
\begin{aligned}
C &= E[\text{Search Length}] + \lambda E[\text{File Length}] \\
&= \sum_{i=1}^{n} L_{\alpha_i} \hat{p}_{\alpha_i} + \lambda \sum_{i=1}^{n} L_{\alpha_i} p_{\alpha_i},
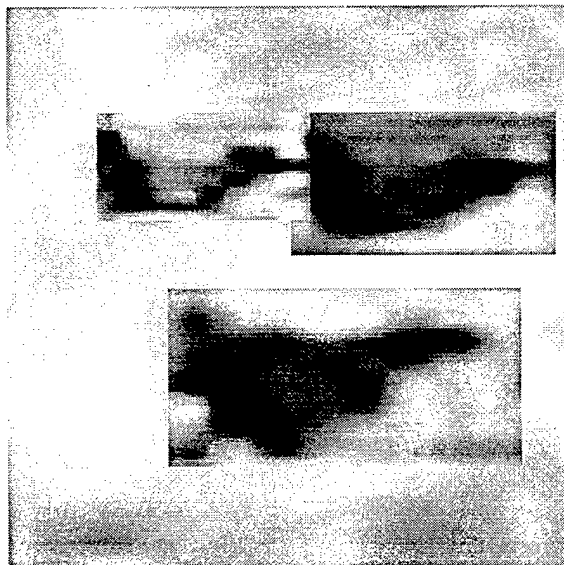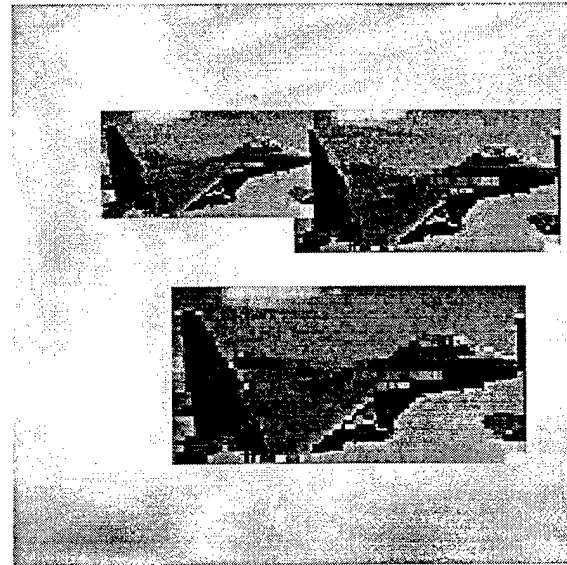\end{aligned}
\tag{3.A.1}
$$

35
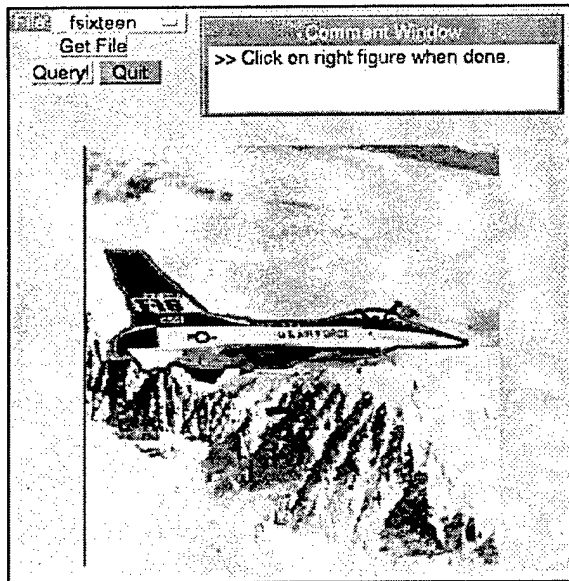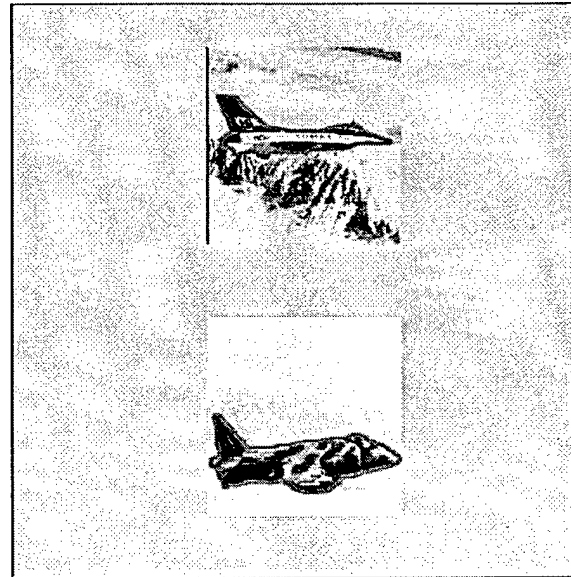
Figure 3.11: Retrieval example, (a) image with query object (plane), and (b) two best retrieved objects.

Figure 3.12: Possible retrieval displays (a) using coarsest subbands and (b) using plane prototype.

where the weights

$$\hat{p}_{\alpha_i} = (p_{\alpha_i} - \sum_{j=1}^{i-1} q_{\alpha_j} p_{\alpha_i \alpha_j}) \tag{3.A.2}$$

are functions of the ordering $\Phi(t)$. Note that search length and file length are measured in bits. The second summand in Eq. 3.A.1 is *independent* of the term ordering and query probabilities.

To minimize C, we need to: 1) define an order $\Phi(t)$ for the terms in the file header, and 2) assign a codeword of length L_{t_i} to each term t_i.

First we describe how to obtain the ordering $\Phi(t)$ which minimizes the expected search length (i.e., minimizing C with $\lambda = 0$). Here we assume that codewords have already been assigned to the terms using a simple Huffman coder. The ordering $\Phi(t)$ of the term set $\{t_1, t_2, \ldots, t_n\}$ which minimizes the expected search length is obtained using the following result. Suppose we have two term orders $\Phi(t)$ and $\Phi'(t)$ which define the order the terms in the header in an identical fashion except for *two* adjacent elements, i.e.,

$$\Phi(t): \quad \ldots t_i t_j \ldots$$
$$\Phi'(t): \quad \ldots t_j t_i \ldots \tag{3.A.3}$$

Let $E[\text{SL}]$ and $E[\text{SL}]'$ denote their expected search lengths. Then $E[\text{SL}] < E[\text{SL}]'$ if and only if

$$\frac{L_{t_i}}{q_{t_i}} < \frac{L_{t_j}}{q_{t_j}}. \tag{3.A.4}$$

This result tells us that, given the codewords, the criteria for minimizing search length is completely independent of the probabilities p_{α_i}. This is a generalized result of [45] which deals with merging files in a storage device to minimize seek time. In that work, the files are known to exist, i.e., $p_{\alpha_i} = 1$ for all i. In our case, p_{α_i} is arbitrary.

Using this result, we may easily order the terms to minimize the expected search length. Simply define $\Phi(t)$ in an arbitrary manner to initialize the ordering. Compare α_{n-1} with α_n using Eq. 3.A.4, swapping the order of the two terms if necessary. Then compare α_{n-2} with α_{n-1}, and so on. After n compares, the correct t_i will be assigned to α_1. The procedure is repeated by comparing α_{n-1} with α_n and comparing adjacent terms until we obtain α_2, etc. The resulting algorithm relies on simple comparisons and can be performed in $O(n^2)$.

The result in Eq. 3.A.4 is intuitively pleasing. For example, if we assume all queries are equiprobable, then the file header is obtained simply by concatenating the codewords of all the terms which appear in that file, shortest codeword length first. Thus, to search for a particular query term in a file collection, we search each file header sequentially for the corresponding term codeword. The search is terminated when we find the codeword (successful) or when we find a codeword whose length is greater than the length of the codeword we are searching for (unsuccessful). Generally, for each query term t_j in the file collection, we have a ratio L_{t_j}/q_{t_j}. To search the collection for term t_i, each file header is searched sequentially. For each term t_j read, the ratio L_{t_j}/q_{t_j} is obtained from a look-up table and compared with L_{t_i}/q_{t_i}. If $L_{t_i}/q_{t_i} < L_{t_j}/q_{t_j}$ at any point in the file header, we terminate the search early as the query term does not exist in that file.

To minimize search length, the codeword lengths L_{α_i} must be known. We have developed a procedure to simultaneously minimize both expected query length and file size as expressed in Eq. 3.A.1 which can be rewritten as

$$C = \sum_{i=1}^{n} L_{\alpha_i} \left((1+\lambda)p_{\alpha_i} - \sum_{j=1}^{i-1} q_{\alpha_j} p_{\alpha_i \alpha_j}\right) = \sum_{i=1}^{n} L_{\alpha_i} w_{\alpha_i} \tag{3.A.5}$$

The algorithm is shown in Fig. 3.13.

In practice, it is reasonable to assume that the term probability estimates p_{t_i} remain fixed for a given collection of documents. As a result, the codeword and codeword length assigned to each term remain fixed. With codeword lengths fixed, files remain fixed in size. The file headers can be regularly updated to minimize expected search length by simply reordering the terms in the header according to a new set of query probabilities. This could occur, for example, once a fixed number of queries have been posed. This method is suboptimal as codewords do not change. However, the resulting system remains close to optimal if term probability estimates p_{t_i} do not change drastically.
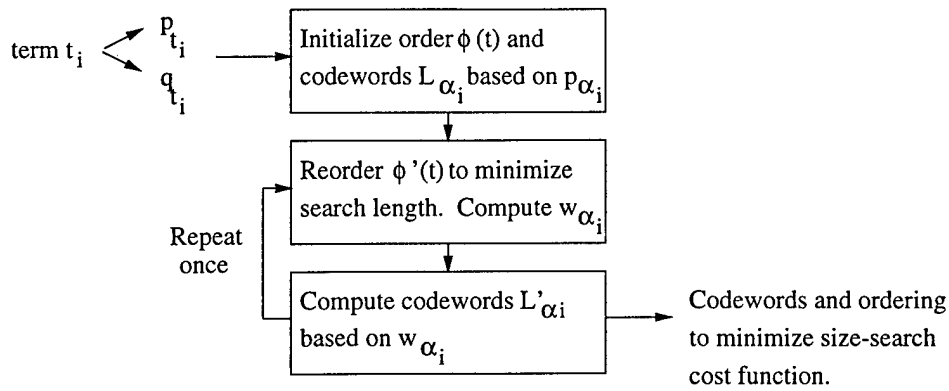
37

Figure 3.13: Given a probabilities of occurrence and of being queried for each term, the algorithm which assigns codewords and ordering to minimize the size-search cost function.
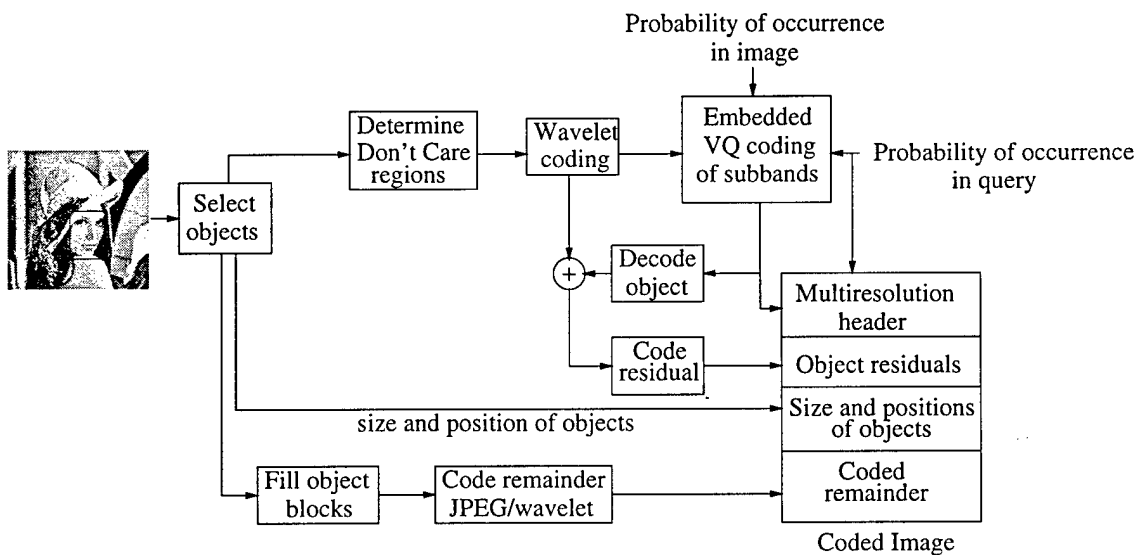
Figure 3.14: Block diagram of image coding algorithm.

Appendix 3.B: Image Coder

The block diagram of our image coder is shown in Fig. 3.14. In the next couple sections, we describe the image coder in detail.

3.B.A   Preprocessing

The first step consists of defining image objects (see, e.g., Fig. 3.5). Objects are first outlined by the user by defining a rectangle about the object with a mouse. The next step is to extract the object from the background. As each object block is defined, a close-up of the object appears in a separate window (c.f. Fig. 3.15). Within this window, the user clicks on several boundary points of the actual object. A gradient-based search algorithm is used to extract the object curve. Once the object curve is defined, the exterior of the curve is filled with black (i.e., zeros) to ensure that background doesn't play a factor when mapping to the VQ dictionaries.

The object blocks are removed from the original image. The remaining image regions are coded using a simple block coding algorithm (e.g., JPEG [33]). The JPEG codewords are stored in the last part of the encoded file (c.f. Fig. 3.14). In addition, the position of each object block is stored for decoding purposes.

38

Figure 3.15: Extracting object from background (a) defining object boundaries, and (b) final object.
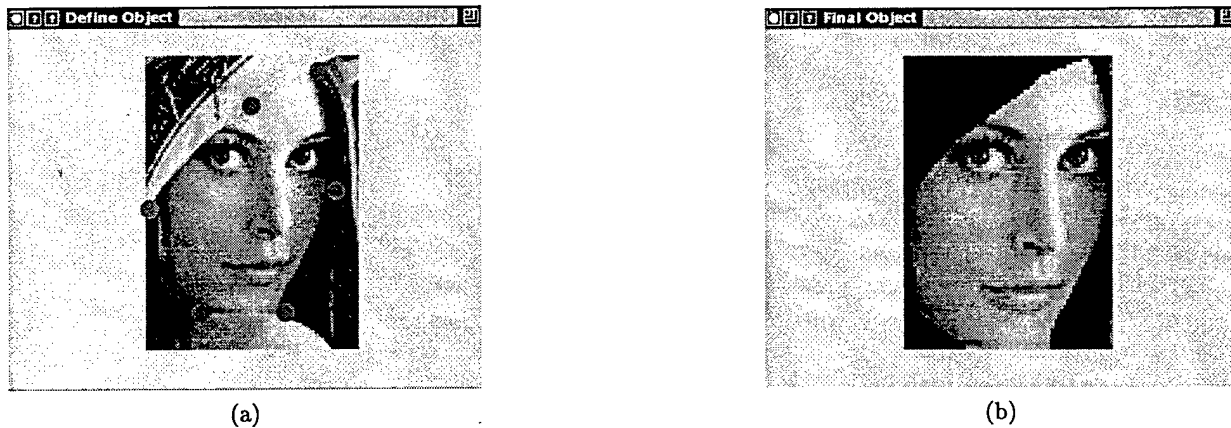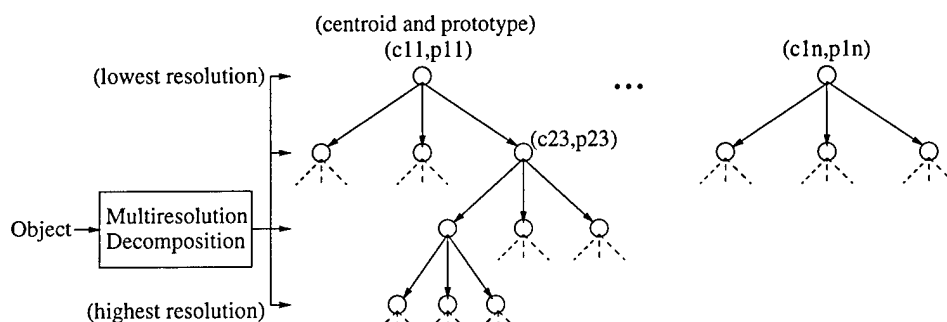

Figure 3.16: Structure of embedded VQ codebooks and prototypes.

The next step consists of applying a wavelet transform [43] to each object block. The number of levels in the wavelet decomposition depends on the original size of the object block. Specifically, the wavelet transform is applied until the lowest frequency subband is 8×8. In this way, the powerful multiresolution characteristic of a wavelet decomposition is exploited for recognition (c.f. Fig. 3.6).

3.B.B Object coding and embedded VQ

Once objects are defined and transformed as above, our new coding technique is applied. The coder constructs the file header in terms of the wavelet transform subbands of the segmented objects to minimize the size-search cost function (Eq. 3.1). To obtain the probability and query probability estimates used by our coding algorithm, we use vector quantization (VQ) [44] to map the subbands to finite dictionaries.

The structure is shown in Fig. 3.16. The VQ dictionaries constructed with our algorithm are *embedded* by scale. The initial VQ dictionary consists of n terms (centroids c_{1i}) based on the lowest resolution subbands of the image objects. Each of the n terms in this dictionary, in turn, has an associated dictionary whose terms are based on the next higher resolution subbands. Since the finite dictionaries are embedded from one scale to the next scale, the probability and query probability assigned to each dictionary entry subband depend on previous (coarser) subbands, i.e., $p_{\alpha_{ij}} = p_{\alpha_{ij}|\alpha_{(i-1)k}} p_{\alpha_{(i-1)k}}$. The probabilities associated with each term in the dictionary determine the proper location and codeword length of each subband codeword in the compressed file header in the exact same manner as in the text case.

To help take into account visual similarity, the mapping of an object's subbands onto the embedded VQ dictionaries is based on a perceptual distance measure. In particular, we construct a frequency domain masking model of each object [46]. The masking model determines when a signal component is perceptually invisible in the presence of another (masking) signal component. When computing distances to obtain the appropriate VQ dictionary entry, we set all masked components of the error image to zero, i.e., only non-masked components contribute to the norm of the error.
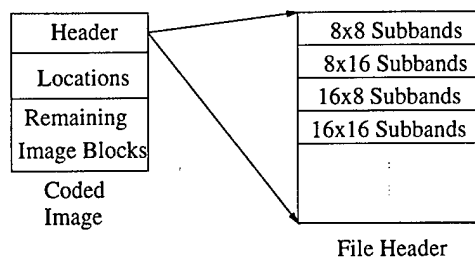
39

Figure 3.17: Ordering of file header in coded image.

A prototype object is also defined for each VQ codebook term. In addition to having a centroid c_{ij}, each entry in the VQ codebooks has an associated prototype object p_{ij}. The prototype object for a codebook entry is defined as the object whose subband at that resolution (scale i) is nearest to the centroid c_{ij}. Note that the prototypes are *actual objects* contained in the image collection. For example, consider the first terms in the lowest resolution codebook. Every object in the image collection whose coarsest subband is mapped to centroid c_{11} is a candidate prototype. The object whose coarsest subband is nearest to c_{11} is used as its prototype, p_{11}. The prototype object is used to represent all of the other objects near it at a given scale. At low scale, the number of objects mapped to a codeword is large and the prototype is "low" in resolution. At higher scale, the prototype becomes more refined as the region it represents diminishes.

To increase retrieval speed, the image file header is ordered in a coarse-to-fine manner. As shown in Fig. 3.17, the first layer of the file header consists of the lowest resolution (coarsest) subband. Within this layer, the coarsest subband components (8×8 subbands) of all the objects in the image are ordered according to probability estimates obtained from the VQ mapping and codeword lengths. The next layer in the file header consists of all the 8×16 subbands of the objects in the image. The remaining subbands are likewise ordered in the file header. Also, the object residual for each subband is stored (second part of "Coded image" in Fig. 3.14). The residual is the difference between the dictionary term which is used to represent each subband and the actual subband values. This is efficiently stored to maintain high fidelity between the original and coded image.

40

Chapter 4

Binary Wavelet Transforms

4.1 Introduction

The multiresolution nature of a wavelet decomposition provides signal specific information localized in space or frequency which can be exploited for signal analysis and processing. Wavelet theory has received a great deal of attention recently as seen by the explosion of the literature in this field. Wavelets have been applied to many applications including edge detection [47, 48], image coding [49, 50], filtering [51, 52], stochastic processes and fractal models [53, 54, 55], time-frequency analysis [56, 57], radar [58, 59], etc. The special issues on wavelets of *IEEE Transactions on Information Theory* (March, 1992) and the *IEEE Transactions on Signal Processing* (December, 1993) include several others.

The development of wavelet decomposition theory has occurred almost exclusively for real- and complex-valued functions. The data to be analyzed and the families of basis functions are real- or complex-valued, and arithmetic is performed in the real or complex field. Just as wavelet theory over real and complex fields has proven useful in many applications, wavelet theory over other arithmetic fields holds a great deal of potential.

A field is any arithmetic system in which one can add, subtract, multiply, and divide under the usual properties of associativity, distributivity, and commutativity. Along with the real and complex fields, denoted by $\mathcal{R}$ and $\mathcal{C}$, we have the finite (Galois) fields denoted by $GF(p)$, where p is the characteristic of the field. Whenever p is a prime, we can construct the field $GF(p)$ as the set of elements $\{0, 1, 2, \ldots, p-1\}$ together with modulo-p addition and modulo-p multiplication. The goal of this work is to generalize wavelet theory over the real and complex fields to $GF(2)$, the binary field.

There have been several attempts to generalize wavelet decomposition and perfect reconstruction filter banks to finite fields. The difficulties associated with the design of perfect reconstruction filter banks in $GF(2)$ are discussed in a pioneering work [60] and are addressed further in [61]. It is shown in those papers that there is no universal factorization of paraunitary matrices in $GF(2)$. The lack of universal factorization in $GF(2)$ is unlike the real field case where it is known that any causal FIR paraunitary matrix of degree N can be factored into a product of N degree-one causal paraunitary systems [62]. In contrast to the real case, a complete characterization of binary paraunitary systems is not possible. Note also that several authors have extended wavelet theory to finite fields with characteristics other than 2, e.g., [63] and [64]. However, these constructions exclude the binary field as they rely on Fourier transform techniques that are difficult to use in $GF(2)$. The complications arising with Fourier techniques over $GF(2)$ are addressed in Section 4.2.B .1.

As we shall see in the sequel, the binary wavelet transform that we construct here shares many of the important characteristics of the real wavelet transform. In particular, it yields an output that is similar to a thresholded real field wavelet transform of the underlying binary image (c.f. Section 4.4 and Figs. 4.8 and 4.9). Therefore, our binary wavelet transform of binary images can be used as an alternative to the real-valued wavelet transform of these images in binary image processing applications (e.g., coding, edge detection, recognition, etc.). Furthermore, this binary transform has several distinct advantages over the real transform when applied to binary (e.g., text, facsimiles, fingerprints) or thresholded data. First, the intermediate and transformed data produced by the binary wavelet transform are binary. No quantization

effects are introduced and the decomposition is completely invertible. The entire decomposition is performed in $GF(2)$. Second, it is extremely fast. Modulo-2 arithmetic is equivalent to exclusive-OR operations. Hence, the transform can be performed using simple Boolean operations. Finally, as the data remains in $GF(2)$, operations on the transformed data tend to be much simpler. For example, if the binary wavelet transform is mapped to a classification tree to allow further image analysis and understanding, all decisions in the tree would be based on simple mod-2 additions and multiplications of binary data.

We construct wavelet decompositions in $GF(2)$ by introducing a new binary field transform which characterizes the filtering properties of binary signals. The new transform allows us to avoid both paraunitary polyphase matrix characterizations and Fourier transform techniques. Furthermore, it retains the concept of spectral classification (e.g., lowpass, highpass, etc.) commonly associated with the Fourier transform. Using the new transform, we construct 1D binary wavelets based on 2-band perfect reconstruction filter banks in $GF(2)$. In the extension to the 2D case, we restrict our attention to separable 2D binary wavelets, where the 2D wavelets are defined as tensor products of 1D wavelets. The theory described here is developed for the analysis of finite images.

We conclude the chapter with an illustration of the potential application of this theory to lossless image coding. In particular, we discuss in Section 4.4 the results of several coding experiments that we have performed. In these experiments, we coded several binary images and their corresponding binary wavelet transform subbands for comparison. No special coding technique (like a modified version of zerotrees [50]) was used on the subbands. Rather, we applied simple run-length coding. The test images included letters (c.f. Figs. 4.11 and 4.13) and fingerprints (c.f. Fig. 4.10). Our results indicate that the transform gives a compact representation of each image. In addition, run-length results produced savings in storage requirements ranging from 14% to 70% over similarly coded original images. A coder modified to take advantage of the subband structure should result in even greater savings.

4.2   The Binary Field

This work unfolds in $GF(2)$ where sequence elements take the values '0' and '1', and arithmetic is performed modulo-2. In this section, we review certain properties of this field and construct a new binary field transform which characterizes the response of binary filters.

4.2.A   Linear Algebra and the Binary Field

The multiresolution decomposition of a sequence in terms of wavelet bases takes the form of a linear system. While certain properties of matrices over the binary field differ from properties of matrices over the real and complex fields, matrix arithmetic (i.e., multiplication, addition, inversion, and determinant calculation) is performed in $GF(2)$ as it is in $\mathcal{R}$ and $\mathcal{C}$. The only difference is that all computations are followed by a modulo-2 operation.

We will use bold faced capital letters (e.g., $\mathbf{A}$) and bold faced small letters (e.g., $\mathbf{v}$) to denote matrices and vectors, respectively. In addition, we will denote by $A(j,k)$ the (j,k) element of $\mathbf{A}$, and by $v(k)$ the k element of $\mathbf{v}$. $\mathbf{A}(:,j)$ and $\mathbf{A}(i,:)$ are the jth column and i^{th} row of $\mathbf{A}$, respectively. $\mathbf{A}^T$ is the transpose of $\mathbf{A}$. $\mathbf{A}^{-1}$ is the inverse of a square matrix $\mathbf{A}$. All indices start at 0.

Matrices and vectors over $GF(2)$ have some special properties that will be important in this chapter [60]. For example, note that for a vector $\mathbf{v} \in \mathcal{R}$ or $\mathcal{C}$, $\mathbf{v}^T\mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$. On the other hand, a vector in $GF(2)$ with an even number $2n$ of nonzero entries has zero energy (since $0 \equiv 2n \bmod 2$). In other words, the energy of a nonzero vector in $GF(2)$ may be zero. Note also that since the only nonzero element in $GF(2)$ is 1, an $N \times N$ matrix $\mathbf{A}$ in $GF(2)$ is invertible if and only if $\det(\mathbf{A}) = 1$. Finally, no column of a unitary matrix $\mathbf{U}$ in $GF(2)$ may have all elements equal to 1. An $M \times N$ matrix $\mathbf{U}$ is unitary in $GF(2)$ if $\mathbf{U}^T\mathbf{U} = \mathbf{I}_N$. In other words, a matrix is unitary only if the columns have unit energy and each pair of columns is mutually orthogonal. If M is even, no column of $\mathbf{U}$ can have all its elements equal to 1 since such a column would have zero energy. If M is odd, no column of $\mathbf{U}$ can have all its elements equal to 1 since all the other columns must then have an even number of 1's by the orthogonality condition. This, however, implies that the other columns have zero energy. An example of a 4×4 unitary matrix $\mathbf{U}$ in $GF(2)$ is shown

below

$$\mathbf{U}^T\mathbf{U} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} = \mathbf{I}_{4\times4}.$$

We also consider circulant matrices. An $N \times N$ matrix $\mathbf{C} = [c_{jk}]_{j,k=0,...,N-1}$ with entries from some field $\mathcal{F}$ is called circulant if its rows are generated by successive shifts of the first row in the matrix. A shift by 1 generates a *one*-circulant matrix. As a *one*-circulant matrix $\mathbf{C}$ is specified by its first row, we will often denote $\mathbf{C}$ by $1 - circ(c_0, c_1, \ldots, c_{N-1})$. A length N circular convolution of two sequences leads naturally to a representation of one sequence as an $N \times N$ *one*-circulant matrix and the other sequence as a length N vector.

Sequence decimation is an inherent operation of filter banks and leads to the second circulant form we are interested in. Decimation of a sequence by a factor m leads to a new sequence which is composed of every m^{th} sample of the original sequence. Suppose that we circularly convolve two sequences of even length N and follow the convolution operation with decimation by a factor of two. Then the equivalent *one*-circulant matrix followed by decimation can be replaced with an equivalent $N/2 \times N$ *two*-circulant matrix of the form

$$\mathbf{D} = \begin{bmatrix} d_0 & d_1 & d_2 & \ldots & d_{N-1} \\ d_{N-2} & d_{N-1} & d_0 & \ldots & d_{N-3} \\ d_{N-4} & d_{N-3} & d_{N-2} & \ldots & d_{N-5} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_2 & d_3 & d_4 & \ldots & d_1 \end{bmatrix}. \tag{4.1}$$

The matrix in Eq. 4.1 is specified by its first row and will be referred to as $2 - circ(d_0, d_1, \ldots, d_{N-1})$. The binary wavelet decomposition we consider in later sections is based on circular convolution of binary sequences with binary filters (wavelet and scaling function coefficients) followed by decimation by 2.

Let us now present a new theorem that we will use in the sequel to determine if a given filtering operation in $GF(2)$ followed by decimation can be inverted. The proof of this theorem relies on transforming circulant matrices into upper Hessenberg matrices. Fig. 4.1 shows the structure of an upper Hessenberg matrix. It resembles a staircase version of an upper triangular matrix. For N even, an $N \times N$ upper Hessenberg matrix has $N/2$ submatrices of size 2×2 along the main diagonal. A straightforward analysis shows that the determinant of such a matrix in $GF(2)$ is equal to the product of the determinants of the 2×2 matrices along the main diagonal.

Now note that in a 2-band filter bank (c.f. Fig. 4.3), a signal is passed simultaneously through two filters, and the filter outputs are decimated by two. An equivalent filtering operation can be obtained by combining the two $N/2 \times N$ *two*-circulant matrices (Eq. 4.1) into one $N \times N$ matrix. The following theorem determines whether the resulting linear system is invertible.

Theorem 1 *Let* $\mathbf{C} = 2 - circ(c_0, c_1, \ldots, c_{N-1})$ *and* $\mathbf{D} = 2 - circ(d_0, d_1, \ldots, d_{N-1})$ *be* $N/2 \times N$ *two-circulant matrices in* $GF(2)$ *where* $N = 2^k$ *for some* $k \geq 1$. *If we construct an* $N \times N$ *matrix* $\mathbf{T}$ *as*

$$\mathbf{T} = \begin{bmatrix} \mathbf{C} \\ \mathbf{D} \end{bmatrix}, \tag{4.2}$$

then $\mathbf{T}$ *can be put in upper Hessenberg form using a pair of fixed transform matrices that are independent of* $\mathbf{T}$. *Furthermore, the determinant of* $\mathbf{T}$ *takes the form*

$$det(\mathbf{T}) = \sum_{i=0,even}^{N-2} c_i \sum_{i=1,odd}^{N-1} d_i + \sum_{i=1,odd}^{N-1} c_i \sum_{i=0,even}^{N-2} d_i. \tag{4.3}$$

Proof. In Appendix 4.A we construct matrices which transform a $2^k \times 2^k$ *two*-circulant matrix of the form Eq. 4.2 into upper Hessenberg form. The $N/2$ submatrices of size 2×2 along the main diagonal of the Hessenberg matrix are all equal and take the form

$$\begin{bmatrix} \sum_{i=0,even}^{N-2} c_i & \sum_{i=1,odd}^{N-1} c_i \\ \sum_{i=0,even}^{N-2} d_i & \sum_{i=1,odd}^{N-1} d_i \end{bmatrix},$$

which has a determinant given by Eq. 4.3. Since $x^p = x$ for any $x \in GF(2)$ and integer $p \geq 1$, the product of $N/2$ such expressions equals the original expression.

Consider the following example of a general 4×4 matrix $\mathbf{T}$ as described in Theorem 1. Using the results of Appendix 4.A, we can form an upper Hessenberg matrix as

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_2 & c_3 & c_0 & c_1 \\ d_0 & d_1 & d_2 & d_3 \\ d_2 & d_3 & d_0 & d_1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} c_0 + c_2 & c_1 + c_3 & c_2 & c_3 \\ 0 & 0 & c_0 + c_2 & c_1 + c_3 \\ d_0 + d_2 & d_1 + d_3 & d_2 & d_3 \\ 0 & 0 & d_0 + d_2 & d_1 + d_3 \end{bmatrix}.$$

After applying a simple permutation, we have the upper Hessenberg form. The determinant of $\mathbf{H}$ can be calculated as

$$((c_0 + c_2)(d_1 + d_3) + (c_1 + c_3)(d_0 + d_2))^2 = (c_0 + c_2)(d_1 + d_3) + (c_1 + c_3)(d_0 + d_2)$$

which agrees with Eq. 4.3.

4.2.B Filtering in the Binary Field

The filtering properties of a sequence over the real field are typically obtained through the application of the discrete Fourier transform (DFT). These properties are used to characterize the wavelet and scaling coefficients of the real field wavelet transform. We will see that as the length N of a DFT over $GF(2)$ increases, the computational complexity of the DFT becomes unmanageable. This leads us to introduce a new binary field transform.

4.2.B .1 Discrete Fourier Transform over $GF(2)$

Although the DFT is usually associated with the real and complex sequences, it is not defined exclusively over these fields. A more general definition is as follows [65]. Let $\mathbf{v}$ be a length N vector over an arbitrary field $\mathcal{F}$. The DFT of $\mathbf{v}$ is another vector $\tilde{\mathbf{v}}$ over $\mathcal{F}$ whose elements are given by

$$\tilde{v}_k = \sum_{i=0}^{N-1} \omega^{ik} v_i, \qquad k = 0, \dots, N-1, \tag{4.4}$$

where ω is an element of order N in the field $\mathcal{F}$. An element of order N satisfies $\omega^N = 1$ and $\omega^k \neq 1$ for $k = 1, 2, \dots, N-1$. The computation of the DFT is performed in the field $\mathcal{F}$ (i.e., modulo-p if $\mathcal{F}$ is $GF(p)$) and all properties typically associated with the Fourier transform hold. Furthermore, if $\tilde{\mathbf{v}}$ is the DFT of a vector $\mathbf{v}$, then $\mathbf{v}$ can be recovered from $\tilde{\mathbf{v}}$ by the inverse DFT. While the DFT is very useful, certain fields may have no element of order N. Therefore, it is not possible to evaluate an N point DFT in these fields. For example, $GF(2)$ has only one non-zero element $\omega = 1$. That element is of order 1. Hence, *we can only compute the DFT of a length 1 sequence in $GF(2)$*.

When a field $\mathcal{F}$ has no element of order N, it is still sometimes possible to obtain a length N DFT by entering an extension field $\mathcal{F}^m$ of $\mathcal{F}$. An extension field is created by using a construction called the polynomial representation of the extension field [65]. The technique involves finding a polynomial $p(x)$ of degree m which is irreducible over $\mathcal{F}$. All elements of $\mathcal{F}^m$ are polynomials of degree at most $m-1$ with coefficients taken from $\mathcal{F}$. Addition in this new field is defined as polynomial addition and multiplication is defined as polynomial multiplication modulo the polynomial $p(x)$.

Using this technique, $GF(2)$ can be extended to $GF(2^m)$ using appropriate polynomials of degree m. However, for a given length N, there may be no extension field in which one may compute a length N DFT. In particular, one cannot compute an *even* length N DFT for any m [65]. Furthermore, for $GF(2^m)$, the highest order of any element is $2^m - 1$. Therefore, we have to extend $GF(2)$ to an arbitrarily large field to compute an arbitrarily long DFT. For example, to compute a length $N = 2^m - 1$ DFT, we need to increase our data

44

dimension from 1 to m since we have to use polynomials of degree m. In contrast, computing the DFT of a real-valued signal increases the data dimension from 1 to 2 (real and imaginary parts) independent of length N. Working with high degree polynomials in $GF(2)$ makes transform computation and interpretation very difficult. To avoid the increase in complexity associated with introducing extension fields, we will use a new binary field transform.

Note that the difficulties associated with the DFT in $GF(2)$ also arise with other transforms related to the DFT. For example, the discrete Hartley transform over $GF(2)$ exists only for lengths N for which the DFT is defined [66]. Therefore, it exists only for lengths N that divide $2^m - 1$ for some m. In particular, N cannot be even. As even length convolutions are essential for 2-band filter banks (due to decimation), the Hartley transform over $GF(2)$ cannot be used to characterize filters in a 2-band filter bank setting.

4.2.B .2   Binary Field Transform

Let us now return to the real field and review an alternative transform to the DFT which employs a class of nonsinusoidal orthogonal waveforms as its set of basis functions. The Walsh-Hadamard transform (WHT) is a real field transformation which has found applications in several signal processing and communications areas [67, 68]. The WHT is based on a complete orthonormal set of rectangular functions known as Walsh functions. Walsh functions take values '-1' and '1' and form a closed set [68]. The first four Walsh functions labeled (a)-(d) are shown in Fig. 4.2.

What makes the individual Walsh waveforms unique is the frequency at which transitions from '-1' to '1' and from '1' to '-1' occur. This frequency, referred to as *sequency*, is defined as one-half the average number of zero-crossings per unit time [67]. Sequency can be thought of as a generalization of frequency as it can be applied to functions whose zero-crossings may occur at irregular intervals and which may be aperiodic. For the example in Fig. 4.2(a)-(d), the waveform sequencies are 0, 1, 1 and 2, respectively (by convention, the waveforms in (b) and (d) are extended to include a final transition near 1). The definition of sequency can be modified to include discrete functions. Specifically, if the number of *sign changes* of a discrete function equals k, the sequency of the discrete function is defined as $k/2$ when k is even and $(k+1)/2$ when k is odd.

By extending on the ideas described above, we construct the binary field transform (BFT) for sequences in $GF(2)$. The basis vectors of the BFT are rectangular waveforms which take values '1' and '0' with varying sequencies. In this case, we define sequency of a binary sequence in terms of transitions from '0' to '1' and '1' to '0' rather than sign changes. These basis vectors constitute the columns of our binary transform matrix. Note that a Fourier series expansion of a length N periodic sequence includes sinusoidal terms up to a frequency of $N/2$. Likewise, our binary field transformation of a length N periodic sequence includes rectangular waveform components with sequencies up to $N/2$. Through the use of this transformation, we replace the conventional frequency information associated with the Fourier transform with sequency information obtained from the binary transform.

For finite sequences, the BFT takes the form of a square symmetric matrix over $GF(2)$. Since the BFT must be invertible, the basis vectors of the BFT *cannot* be Walsh functions with -1's replaced by 0's. The determinant of the matrix corresponding to a WHT is even, i.e., it is zero in $GF(2)$. Hence, such a matrix is non-invertible in $GF(2)$. We therefore proceed to construct the $N \times N$ BFT matrices $\mathbf{B}_N$ recursively as follows. We define

$$\mathbf{B}_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \tag{4.5}$$

and

$$\mathbf{B}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}. \tag{4.6}$$

For $N \geq 6$ and even, we construct $\mathbf{B}_N$ in terms of four submatrices as

$$\mathbf{B}_N = \begin{bmatrix} \mathbf{B}_N^{ul} & \mathbf{B}_N^{ur} \\ \mathbf{B}_N^{ll} & \mathbf{B}_N^{lr} \end{bmatrix}. \tag{4.7}$$

The upper-left $(N-2) \times (N-2)$ submatrix $\mathbf{B}_N^{ul}$ is defined as

$$\mathbf{B}_N^{ul} = \begin{bmatrix} \mathbf{1s}_{2\times2} & \mathbf{1s}_{2\times(N-4)} \\ \mathbf{1s}_{(N-4)\times2} & \overline{\mathbf{B}}_{(N-4)} \end{bmatrix}. \tag{4.8}$$

In the above equation, the subscripts denote the sizes of the submatrices. The matrix $\mathbf{1s}_{N\times M}$ is an $N \times M$ matrix that consists of 1's. Matrix $\overline{\mathbf{B}}_{N-4}$ is the result of applying the logical-NOT operation to each element of the BFT matrix $\mathbf{B}_{N-4}$. The upper-right $(N-2) \times 2$ submatrix $\mathbf{B}_N^{ur}$ and lower-left $2 \times (N-2)$ submatrix $\mathbf{B}_N^{ll}$ are defined as

$$\mathbf{B}_N^{ur} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \tag{4.9}$$

and

$$\mathbf{B}_N^{ll} = \mathbf{B}_N^{urT}. \tag{4.10}$$

Finally, the lower-right 2×2 submatrix $\mathbf{B}_N^{lr}$ is defined as

$$\mathbf{B}_N^{lr} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}. \tag{4.11}$$

For N odd, we construct $\mathbf{B}_{N+1}$ and define $\mathbf{B}_N$ as the upper-left $N \times N$ submatrix $\mathbf{B}(0:N-1, 0:N-1)$ of $\mathbf{B}_{N+1}$.

Note that the BFT matrix takes an analogous form to the Walsh-ordered, or sequency ordered, Walsh-Hadamard transform matrix. That is, the sequency ordering of the columns corresponds to a nondecreasing sequency ordering of Walsh functions. Furthermore, the number of 1's in each of the first $N/2$ columns is even. As an example, consider the 8×8 BFT matrix

$$\mathbf{B}_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}. \tag{4.12}$$

From left to right, the sequency value of each column is 0, 1, 1, 2, 2, 3, 3, and 4.

Note that the first column of a BFT matrix corresponds to a DC basis vector of zero sequency. Every element in that first column is equal to 1. Therefore, the BFT matrix $\mathbf{B}_N$ *cannot* be orthonormal. Fortunately, each $\mathbf{B}_i$ is invertible over the binary field (i.e., $\det(\mathbf{B}_i) = 1$ for all $i \geq 2$) and its inverse $\mathbf{B}_N^{-1}$ can be evaluated using a simple recursive formula. In particular, we show in Appendix 4.B that the inverse of $\mathbf{B}_N$ for $N \geq 6$ is given by

$$\mathbf{B}_N^{-1} = \begin{bmatrix} \mathbf{A}_{(N-2)\times(N-2)} & \mathbf{C}_{(N-2)\times2} \\ \mathbf{C}_{(N-2)\times2}^T & \mathbf{D}_{2\times2} \end{bmatrix}. \tag{4.13}$$

The sizes of the submatrices are given by their subscripts. They are constructed as

$$\mathbf{A}_{(N-2)\times(N-2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 1 & 1 & 0 & \cdots & 0 & 1 \\ 0 & 1 & & & & & \\ 0 & 0 & & \mathbf{B}_{N-4}^{-1} & & & \\ \vdots & \vdots & & & & & \\ 0 & 0 & & & & & \\ 1 & 1 & & & & & \end{bmatrix}, \tag{4.14}$$

$$\mathbf{C}_{(N-2)\times 2} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \qquad (4.15)$$

and

$$\mathbf{D}_{2\times 2} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \qquad (4.16)$$

For the 8×8 BFT example, the inverse matrix is

$$\mathbf{B}_8^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \qquad (4.17)$$

Hence, the construction of both transform and inverse transform matrices relies solely upon shifting and embedding matrix elements. *No* matrix computation is required.

Each BFT matrix of size N forms a basis for vectors of length N in $GF(2)^N$. Specifically, every vector in the space is a unique combination of the vectors which compose the columns of the BFT matrix. The columns of the matrix, in turn, are associated with the specific sequences defined above. As a result, vectors in $GF(2)$ can be written *uniquely* in terms of different sequency components ranging in value from 0 to $N/2$.

We compute the binary field transform in a manner that is different from that commonly associated with DFT or WHT spectra. Rather than computing the matrix-vector product $\mathbf{B}_N^{-1}\mathbf{x}$ for a length N sequence $\mathbf{x}$, we apply $\mathbf{B}_N^{-1}$ to all circular shifts of $\mathbf{x}$. Specifically, to compute the $N \times N$ matrix BFT $\hat{\mathbf{X}}$ of $\mathbf{x}$, we begin by forming the equivalent *one*-circulant matrix $\mathbf{X} = 1 - circ(\mathbf{x})$. We then evaluate the matrix-matrix product

$$\hat{\mathbf{X}} = \mathbf{X}\mathbf{B}_N^{-1}. \qquad (4.18)$$

This definition of spectrum in $GF(2)$ is motivated by the fact that a circular shift of $\mathbf{x}$ can lead to a different transform. In particular, a shift by one time or space unit in the discrete time or space domain is not equivalent to multiplication of the transform by a simple factor as with the Fourier transform on the real field. Another advantage of this definition is that it allows us to define the magnitude of the spectral components of $\mathbf{x}$ along each BFT basis vector simply by counting the number of 1's in the corresponding column of $\hat{\mathbf{X}}$.

The BFT as defined above has three distinct advantages over the generalized DFT discussed in Section 4.2.B .1. First, no *polynomial* operations are required to compute the BFT. This significantly reduces computational complexity, as all operations are performed with simple modulo-2 addition. Second, we are able to compute the BFT for any size N input vector. The DFT, on the other hand, could only be computed for certain size N (excluding, for example, any even length N). Third, the BFT matrix is easier to interpret than a multidimensional DFT.

A limitation of the BFT, as we have defined it above, is that the circular convolution of two sequences is *not* equivalent to multiplication of their transforms in the BFT domain. In particular, a filter that has large spectral components at high sequencies need *not* be a high pass filter! To characterize a filter as a lowpass, bandpass or highpass filter, we need to examine its effect on the basis vectors of a BFT. This may be done by evaluating the circular convolution of the filter with each basis vector via the product $\tilde{\mathbf{X}} = \mathbf{X}\mathbf{B}_N$. Note that this is akin to evaluating a BFT with respect to the *inverse* BFT matrix $\mathbf{B}_N^{-1}$. We shall refer to $\tilde{\mathbf{X}}$ as the filter binary field transform (FBFT) of $\mathbf{x}$.

In summary, to study the effect of a filter $\mathbf{h}$ on all circular shifts of a sequence $\mathbf{x}$ we need to examine the FBFT of $\mathbf{h}$ and the BFT of $\mathbf{x}$. In particular, a simple computation shows that the output $\mathbf{y}$ can be computed via the product $\hat{\mathbf{X}}(1,:)\tilde{\mathbf{H}}^T$, where $\hat{\mathbf{X}}(1,:)$ denotes the first row of $\hat{\mathbf{X}}$. Furthermore, the output $\mathbf{y}$

of a circular shift N in $\mathbf{x}$ can be computed via the product $\hat{\mathbf{X}}(N+1,:)\tilde{\mathbf{H}}^T$. While computing the circular convolution of $\mathbf{x}$ and $\mathbf{h}$ in the BFT domain via the product $\hat{\mathbf{X}}(1,:)\tilde{\mathbf{H}}^T$ is not computationally less expensive than computing it directly, we will find in the sequel that the FBFT provides us with a powerful tool for characterizing and designing filters in $GF(2)$.

As an example, consider the vector $\mathbf{x} = [1\ 0\ 0\ 1]$. Its BFT $\hat{\mathbf{X}}$ is given by

$$\hat{\mathbf{X}} = \mathbf{X}\mathbf{B}_4^{-1} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

We interpret this result as follows. Viewed as a signal, $\mathbf{x}$ has no spectral component corresponding to the 1-sequence basis vector $[1011]$ (see $\mathbf{B}_4$ below), a maximum strength spectral component corresponding to the 1-sequence basis vector $[1100]$ and half maximum strength spectral components corresponding to the 0-sequence and 2-sequence basis vectors $[1111]$ and $[1010]$ (DC and high sequency).

On the other hand, its FBFT $\tilde{\mathbf{X}}$ captures its performance as a filter and is given by

$$\tilde{\mathbf{X}} = \mathbf{X}\mathbf{B}_4 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We interpret this result as follows. The FBFT matrix shows that the zero sequency component (column 0) is not passed by the filter. All three other BFT basis vectors are passed by the filter. We conclude that the filter corresponding to the impulse response $\mathbf{x}$ is of a high pass nature. Note in particular that even though the BFT of $\mathbf{x}$ is zero at the location of the 1-sequence basis vector $[1011]$, $\mathbf{x}$ viewed as a filter passes $[1011]$. This of course confirms what we have mentioned above: the circular convolution of two sequences is *not* equivalent to multiplication of their transforms in the BFT domain. By examining $\tilde{\mathbf{X}}$ and $\hat{\mathbf{X}}$ we can predict the effect of filtering $\mathbf{x}$ or circularly shifted versions of $\mathbf{x}$ using a filter with impulse response $\mathbf{x}$. For example, the output of the filter $\mathbf{x}$ driven by the signal $\mathbf{x}$ is equal to the multiplication from the right of the first row of $\hat{\mathbf{X}}$ by $\tilde{\mathbf{X}}^T$. Hence, it is given by $[0101]$. This result is confirmed by direct evaluation of the circular convolution of $\mathbf{x}$ with itself.

4.3 A Theory of 1D Binary Wavelets

The wavelet theory over binary fields that we propose parallels the theory developed over the real field. In particular, we view the construction of 2-band discrete orthonormal binary wavelets as equivalent to the design of a 2-band perfect reconstruction (PR) filter bank with added vanishing moments conditions. The 2-band PR filter bank is shown in Fig. 4.3, where the input signal is simultaneously passed through lowpass L and highpass H filters and then decimated by 2 to give approximation and detail components of the original signal (analysis section). The two decimated signals may then be upsampled and passed through complementary filters and summed to reconstruct our original signal (synthesis section). Often a filter bank is cascaded with 1 or more additional filter banks to provide further resolutions of the input signal. The cascade of filter banks, termed a multiresolution pyramid due to its structure, need not use the same filters at each stage. The conditions which follow can be applied to individual stages in the cascade to ensure that the overall pyramid satisfies the conditions.

To guarantee that we can perform a useful multiresolution decomposition that inherits the important characteristics of the real wavelet decomposition, and still be able to reconstruct our original signal, the filters must satisfy 3 constraints: a bandwidth constraint, a vanishing moments constraint and a perfect reconstruction constraint. Specifically, we restrict the bandwidths of the lowpass and highpass filters to be approximately equal in size. This is needed to guarantee that no information is lost after we downsample the outputs of the two filters by a factor of two. We argue in the sequel that the vanishing moments constraint that real wavelet filters satisfy should be replaced by a constraint on the number of low and high sequency basis vectors that the highpass and lowpass filter block respectively. In particular, this constraint guarantees

that, as in the real field case, the binary wavelet transforms of slowly varying binary sequences are very sparse. Finally, we impose the perfect reconstruction constraint to guarantee that the binary wavelet transform is invertible. We address each of these constraints in the next few sections for the separable 2D case, where the 2D binary wavelets are defined as tensor products of 1D wavelets.

Since each stage in our binary wavelet decomposition involves decimation by a factor of 2, we will assume that the input sequence has a length $N = 2^K$. In particular, all circular convolutions are of length $N = 2^K$ and any filter of even length $N' < N$ will be padded with zeros to length $N = 2^K$. In the sequel, we shall also use l and h to denote the vector representations of the N binary (zero-padded if necessary) scaling and wavelet filter coefficients respectively.

4.3.A Decimated FBFT Computation and Bandwidth

To meet the first condition, we place bandwidth restrictions on the lowpass filter l and highpass filter h. Since we subsample the outputs of the lowpass and highpass filters by 2, we modify the computation of the FBFT discussed in Section 4.2.B .2 to directly incorporate this action. The decimation operation can be taken into account as follows. First, we form the equivalent *two*-circulant matrices (Eq. 4.1) $\mathbf{L}_2 = 2 - circ(\mathbf{l})$ and $\mathbf{H}_2 = 2 - circ(\mathbf{h})$. Next, we compute the decimated FBFT's $\tilde{\mathbf{L}}_2$ and $\tilde{\mathbf{H}}_2$ by retaining every other row of $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{H}}$ respectively, i.e.,

$$\tilde{\mathbf{L}}_2 = \mathbf{L}_2 \mathbf{B}_N = \tilde{\mathbf{L}}(0:2:N-1,:)$$
$$\tilde{\mathbf{H}}_2 = \mathbf{H}_2 \mathbf{B}_N = \tilde{\mathbf{H}}(0:2:N-1,:). \tag{4.19}$$

Note that the resulting FBFT matrices $\tilde{\mathbf{L}}_2$ and $\tilde{\mathbf{H}}_2$ are of size $N/2 \times N$. For the bandwidth condition, we restrict the bandwidths of l and h to be approximately equal in size. Specifically, the number of nonzero columns of $\tilde{\mathbf{L}}_2$ and $\tilde{\mathbf{H}}_2$ are approximately equal. This helps maintain an even distribution of sequency content in the filter outputs. The even distribution allows for a better multiresolution decomposition of the input data, as we are able to successively half the resolution of the input data as we proceed through a multistage decomposition. In addition, we require l to be a lowpass filter, i.e., we set

$$\tilde{\mathbf{L}}_2(:,N-1) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.20}$$

4.3.B Vanishing Moments

The vanishing moments property of wavelets over the real field ensures that the Fourier transform of the wavelet coefficients will have zeros of a certain order at DC, i.e., it will decay smoothly to zero as the frequency approaches zero. In particular, a large number of vanishing moments leaves minimal power in the low frequency region of the Fourier transform of the wavelet. Therefore, a filter corresponding to a wavelet with a larger number of vanishing moments has a *better low frequency attenuation* performance (c.f. Fig. 4.4). This guarantees a degree of smoothness in the wavelet in the time domain and allows for the compact representation of slowly varying data signals.

The vanishing moments criteria takes on a somewhat different interpretation in $GF(2)$. As we cannot have a "smooth" decay to zero in $GF(2)$, the frequency domain smooth decay to zero as frequency approaches zero and high attenuation characteristics at low frequencies are replaced with multiple consecutive zeros at low sequency in the FBFT matrix. We will see in Section 4.4 that this property allows for a more compact representation of slowly varying $GF(2)$ sequences. Compactness means here that filtered sequences consist mostly of zeros. This is intuitively apparent. The additional zeros at low sequency allow us to avoid representing (passing) the slowly varying regions in an input sequence. The wavelet filter output mainly represents edges and regions of a varying nature.

Note that in the real field, the vanishing moments property is crucial for pointwise convergence of the continuous time wavelet derived from the discrete-time filter [69]. We have not addressed this issue yet for our binary filters. Here, we are simply addressing binary wavelet theory for discrete-time finite-sized sequences.

Based on the above discussion, our second condition states that the wavelet coefficients do not pass the DC sequence component. In terms of the FBFT,

$$\tilde{\mathbf{H}}_2(:,0) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4.21}$$

The above equation guarantees one vanishing moment. We can define additional vanishing moments by forcing the filter $\mathbf{h}$ to block more low sequence basis components. As in the real case, the resulting wavelet filters would then have better bandpass or highpass characteristics.

To impose two vanishing moments we set the first two columns of $\tilde{\mathbf{H}}_2$ to zero. Now observe that the decimated FBFT is computed by multiplying *two*-circulant matrices (which shift filter coefficients by two in each row) from the right by $\mathbf{B}_N$. Furthermore, the first and second columns of $\mathbf{B}_N$ differ only in their last two entries. These last two entries are both equal to one in the first column and both equal to zero in the second column. Therefore, setting the first two columns of $\tilde{\mathbf{H}}_2$ to zero automatically constrains taps $2n$ and $2n+1$ of $\mathbf{h}$, $h(2n)$ and $h(2n+1)$, to be equal, i.e., $h(2n) = h(2n+1)$ for $0 \le n \le N/2 - 1$. Note also that each of the first $N/2$ columns of $\mathbf{B}_N$ has an even number of 1's. The 1's appear in pairs along each column. Specifically, denote by $B_N(n,j)$ the (n,j) element of matrix $\mathbf{B}_N$. If $B_N(2n,j) = 1$ for a given $n \in [0, N/2-1]$ and any $j \in [0, N/2-1]$[1], then $B_N(2n+1,j) = 1$. Therefore, the condition $h(2n) = h(2n+1)$ for $0 \le n \le N/2 - 1$ also forces the first $N/2$ columns of the FBFT matrix to be zero.

Furthermore, recall that, unlike the real field case, N here is the *convolution length* rather than the filter length. As mentioned at the beginning of this section, the convolution length may be equal to or larger (when zero-padding is used) than the number of taps in the wavelet filter. Therefore, the *number of vanishing moments depends on the convolution length rather than the number of filter taps*. Since padding with zeros does not change the fact that $h(2n) = h(2n + 1)$, a filter of even length N' with $N'/2$ vanishing moments will have exactly $N/2$ vanishing moments if it is padded with zeros to length N.

Note also that we can set no more than $N/2$ columns of the FBFT matrix corresponding to the highpass filter equal to zero. If we try to do so, we end up with an overdetermined and inconsistent set of equations. This, in some sense, is the counterpart of the fact that in the real case, an N taps highpass filter corresponding to a wavelet decomposition can have no more than $N/2$ zeros at the origin of the frequency axis. In summary, *the filter $\mathbf{h}$ can only have one or $N/2$ vanishing moments.*

Let us illustrate the points that we have made above with a simple example. Suppose that we wish to design a length $N' = 4$ wavelet filter with $p = 2$ vanishing moments. We compute the decimated FBFT of the 4 tap filter $\mathbf{h}$ as

$$\mathbf{H}_2\mathbf{B}_4 = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 \\ h_2 & h_3 & h_0 & h_1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \sum_i h_i & h_0 + h_1 & h_0 + h_2 + h_3 & h_0 + h_2 \\ \sum_i h_i & h_2 + h_3 & h_0 + h_1 + h_2 & h_0 + h_2 \end{bmatrix}$$

By setting the first two columns to zero we obtain the following equations

$$h_0 + h_1 = 0$$
$$h_2 + h_3 = 0$$

Therefore, the filter $\mathbf{h} = [1\,1\,0\,0]$ will have 2 vanishing moments. If we pad $\mathbf{h}$ with zeros to length N, the resulting filter will have $N/2$ vanishing moments. For example, the FBFT's of $\mathbf{h}$ and $\mathbf{h}$ padded to length 8 are

$$\tilde{\mathbf{H}}_2 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{H}}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

As expected, the first two columns of the original $\tilde{\mathbf{H}}_2$ and the first four columns of the padded $\tilde{\mathbf{H}}_2$ are identically zero. In other words, the filter $\mathbf{h}$ padded to length 8 has 4 vanishing moments.

---

[1] As mentioned in Section 4.2.A, all indices start at zero in our notation.

4.3.C Perfect Reconstruction

For the third condition, we require the binary filters of the multiresolution decomposition to satisfy the perfect reconstruction (PR) property to ensure that the decomposition is invertible. Consider the *two-circulant* matrices $\mathbf{L}_2$ and $\mathbf{H}_2$ of the filter coefficients. When Theorem 1 was introduced, it was noted that we can compute the operation of the decimated filter bank by applying the matrix

$$\mathbf{T} = \left[\begin{array}{c} \mathbf{L}_2 \\ \mathbf{H}_2 \end{array} \right] \tag{4.22}$$

to the input data sequence. The length N output sequence is then ordered with the first $N/2$ values representing the approximation signal (output of the decimator following the lowpass filter) and with the second $N/2$ points representing the detail signal (output of the decimator following the highpass filter). Perfect reconstruction is guaranteed if the matrix $\mathbf{T}$ is invertible. By Theorem 1, we know that $\mathbf{T}$ is invertible if and only if the filter coefficients l_i and h_i satisfy

$$\det(\mathbf{T}) = \sum_{i=0,even}^{N-2} l_i \sum_{i=1,odd}^{N-1} h_i + \sum_{i=1,odd}^{N-1} l_i \sum_{i=0,even}^{N-2} h_i = 1. \tag{4.23}$$

Now observe that Eqs. 4.20 and 4.21, imply that

$$\sum_{i=0,even}^{N-2} l_i = 0$$

$$\sum_{i=0}^{N-1} h_i = 0. \tag{4.24}$$

Therefore, we conclude that we must also have

$$\sum_{i=1,odd}^{N-1} l_i = 1,$$

$$\sum_{i=0,even}^{N-2} h_i = 1, \tag{4.25}$$

$$\sum_{i=1,odd}^{N-1} h_i = 1.$$

The above equations are also equivalent to

$$\tilde{\mathbf{L}}_2(:,0) = \tilde{\mathbf{H}}_2(:,N-1) = \left[\begin{array}{c} 1 \\ \vdots \\ 1 \end{array} \right]. \tag{4.26}$$

4.3.D Filter Design

With the conditions listed above, we can design lowpass and highpass filters for a useful binary wavelet decomposition. As mentioned in Section 4.3.B, the filter design procedure is simplified by the facts that a wavelet filter of length N' can only have one or $N'/2$ vanishing moments and zero-padding retains spectral properties. As a result, we can get a useful decomposition using filters with relatively small support, e.g., two to eight coefficients.

To design the scaling and wavelet filters for a binary wavelet transform of length $N = 2^K$ we proceed as follows. We begin by selecting the length N' of the lowpass and highpass filters. The support (length) of the filters, N', is a design parameter that must be carefully selected depending on the application. The filter support determines how many image coefficients will contribute to each output of the convolution process. Therefore, filters with a small support will produce outputs that depend only on the properties of the image in a small neighborhood of the location of each output sample. We will see in the next section that a useful decomposition can be obtained using relatively short filters. As an example, $h = [1\ 1]$ performed very well in our experiments as it has $N/2$ vanishing moments for a convolution length N. A small support is particularly

51

useful for images composed mainly of high sequency basis vectors since it avoids filtering across multiple edges (e.g., Fig. 4.10). We have used both short and longer filters (e.g., $N = 8$) with images which consisted mainly of low sequency regions. In particular, we have found that with binary 256×256 images it is usually sufficient to use a value for N' that is less than or equal to 8.

Next, we restrict the filter coefficients using Eqs. 4.24 and 4.25. This guarantees PR, a lowpass characteristic for l, and one vanishing moment for h. As Eq. 4.25 reduces the number of degrees of freedom by two, there are $2^{N'-2}$ possible wavelet filters. Finally, we select a highpass filter with a maximum number of vanishing moments. Maximizing the number of vanishing moments reduces the number of degrees of freedom available to design the wavelet filter to $N'/2 - 1$. We have found in our experiments that filters with equal number of vanishing moments and similar support yield equivalent results. This is partially a reflection of the fact that these filters completely block the same set of low sequency basis vectors. Since a filter with $N'/2$ vanishing moments is the best length N' highpass filter that we can design, we select the wavelet filter to be any of the $2^{N'/2-1}$ length N' filters with $N'/2$ vanishing moments. By zero-padding the filter to length $N = 2^K$, we are then guaranteed $N/2$ vanishing moments.

For example, suppose we wish to design a length $N' = 4$ wavelet filter with $p = 2$ vanishing moments. We have seen in Section 4.3.B that the filter coefficients must satisfy

$$h_0 + h_1 = 0$$
$$h_2 + h_3 = 0$$

to guarantee that the filter h has two vanishing moments. The PR condition further implies that

$$h_0 + h_2 = 1$$
$$h_1 + h_3 = 1$$

By combining these two sets of equations, we find that $\mathbf{h} = [h_0 \; h_0 \; (1 + h_0) \; (1 + h_0)]$. Since $h_0 = 0$ or 1, we have two possible wavelet filters with 2 vanishing moments: $[1\,1\,0\,0]$ or $[0\,0\,1\,1]$. Finally, recall from Section 4.3.B that when h is zero-padded to $N > N'$, the number of vanishing moments increases to $N/2$. Therefore, we can use either of these two filters with any length N binary wavelet transform.

The lowpass filter is designed similarly by blocking high sequency BFT basis vectors. Unlike the real field case, the lowpass filter is not completely specified once we have designed the highpass filter. However, its coefficients are still constrained by the PR and bandwidth constraints. For example, we can use either of the lowpass filters $[1\,1\,1\,0]$, $[1\,0\,1\,1]$ with the highpass filter $[1\,1\,0\,0]$. Other choices are also possible when the highpass filter is padded to a length $N \geq 4$.

4.4   2D Binary Wavelets and Examples

Our 2D binary wavelets are tensor products of 1D binary wavelets. In particular, the first stage of a 2D binary wavelet transform of an $N \times N$ image $\mathbf{F}$ involves pre- and post- multiplying $\mathbf{F}$ by $\mathbf{T}$ (Eq. 4.22) and $\mathbf{T}^T$. This corresponds to passing the image through a lowpass 2D separable filter and 3 bandpass 2D separable filters and decimating by 2 in each direction (c.f. Fig. 4.5). The sequency regions of the transformed image are likewise the tensor product of the 1D sequency regions. For example, we can compute a 3-stage decomposition of a binary image using the single stage of Fig. 4.5 by successively applying the decimated output of each LL filter as the input to the next stage. The resulting multiresolution image has sequency regions as shown in Fig. 4.6.

To illustrate the binary wavelet transform, we designed a 2D scaling filter and wavelet with 4 vanishing moments using the procedure developed above. Next we performed a 3-stage multiresolution decomposition of the 256×256 binary plane image shown in Fig. 4.7. The filter coefficients for the lowpass and highpass filters were $\mathbf{l} = [1\;1\;1\;0\;1\;0\;1\;0]^T$ and $\mathbf{h} = [1\;1\;1\;1\;1\;1\;0\;0]^T$, respectively. The result of the 3-stage decomposition is shown in Fig. 4.8. We can clearly see that the higher sequency edge transitions mapped into higher sequency regions of the multiresolution image map. Of the original 65536 pixels, the original image has 47999 nonzero pixels, while the transform has only 2875 nonzero pixels. To measure the compactness of the signal representation, we define an empirical entropy-based cost function

$$H(p) = -p \log_2(p) - (1 - p) \log_2(1 - p)$$

where p is the number of nonzero pixels in the image divided by the total number of pixels in the image. The measure takes values between 0 and 1 and is maximum when an equal number of zero and nonzero pixels exist in the image. A small value indicates a large discrepancy between the number of zero and nonzero coefficients and a more efficient coding representation. This measure corresponds to the Shannon entropy of the image when the pixel values are independent and identically distributed with probability p. Similar cost functions have been used to determine the compactness of different signal representations for best basis selection [70]. For this image, we observe a decrease in entropy from 0.838 bpp to 0.260 bpp. Most of these nonzero pixels are mapped to low sequency regions. A real-valued Daubechies wavelet transform with two vanishing moments was also applied to the binary plane image. The result of applying a threshold operator to the real-valued wavelet decomposition is shown in Fig. 4.9. Transform coefficients > 0 are shown in white. The results are almost identical. Unlike the real-valued wavelet transform which used floating point arithmetic, the binary wavelet transform introduced no quantization errors and was computationally simple.

We also coded several binary images and their corresponding binary wavelet transform subbands using a run-length coder. Test images included letters (c.f. Figs. 4.11 and 4.13) of size 256×256, and fingerprints (c.f. Fig. 4.10) of size 512×512. We applied two levels of decomposition to the letters, and four levels to the fingerprints. The filters used were $\mathbf{l} = [1\ 1\ 1\ 0]^T$ and $\mathbf{h} = [1\ 1\ 0\ 0]^T$. Summaries of the results are shown in Tables 1 and 2. Our results show that entropy was significantly reduced in the transformed images, indicating that we have a more compact representation. In addition, run-length coding results indicate savings in storage sizes ranging from 14% to 70% over similarly coded original images. A coder modified to take advantage of the subband structure should result in even greater savings (e.g., a properly modified zerotree coder).

Wavelet transforms have also been used as a pre-processing tool for character recognition systems [71]. We have applied the binary wavelet decomposition to images of single characters. Two original 256×256 images of white letters on black backgrounds are shown in Figs. 4.11 and 4.13. The one stage decompositions shown in Figs. 4.12 and 4.14 agree well with the expected behavior. In particular, the lowpass images (a) of each letter show a "binary blurring", where the original edges are no longer sharp. The two bandpass images (b) and (c) show the vertical and horizontal edges of the original images, displaying some of the components between approximation and detail. Finally, the highpass image (d) in each case shows that the diagonal edges and corners of the original images were detected and reproduced in the detail images. We clearly see differences in the multiresolution components between the two letters 'a' and 'b'. Additional multiresolution images could be obtained by passing these images through more decomposition stages.

4.5   Conclusion

We have introduced a new sequency-based transform applicable to sequences over $GF(2)$. Based on this transform, a theory of binary wavelets was constructed. It was shown that by constructing the theory in a manner similar to that used in the real field, a PR multiresolution analysis was possible. Using simple modulo-2 operations, the binary wavelet transform yields an output similar to the thresholded output of a real wavelet transform operating on the underlying binary image. Therefore, our binary wavelet transform of binary images can be used as an alternative to the real-valued wavelet transform of these images in binary image processing applications (e.g., coding, edge detection, recognition, etc.). Experimental results indicate character recognition and lossless subband coding are promising areas of applications for this theory.

Appendix 4.A

The derivation of Theorem 1 relied on the fact that *two*-circulant matrices could be transformed into the upper Hessenberg form *independent* of the circulant matrix. The transformation of a *two*-circulant matrix $\mathbf{T}$ into an upper Hessenberg matrix $\mathbf{H}$ is computed as

$$\mathbf{H} = \mathbf{STR} \tag{4.A.1}$$

where $\mathbf{S}$ and $\mathbf{R}$ are the left and right transform matrices. All matrices are of size $N \times N$ where N is even.

Consider the hybrid *two*-circulant matrices $\mathbf{T}$ (Eq. 4.2). A natural starting point to obtain the upper Hessenberg form would be to zero the lower $N/2 - 1$ entries of the first and second columns of the $N/2 \times N$

submatrices $\mathbf{C}$ and $\mathbf{D}$. This is accomplished by adding the even columns, $\mathbf{T}(:,j)$, $j = 2, 4, \ldots, N - 2$, to column 0 and the odd columns, $\mathbf{T}(:,j)$, $j = 3, 5, \ldots, N - 1$, to column 1. The new matrix takes the form

$$\mathbf{T}' = \begin{bmatrix} \sum_{i=0,even}^{N-2} c_i & \sum_{i=1,odd}^{N-1} c_i & c_2 & \cdots & c_{N-1} \\ \sum_{i=0,even}^{N-2} c_i & \sum_{i=1,odd}^{N-1} c_i & c_0 & \cdots & c_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=0,even}^{N-2} c_i & \sum_{i=1,odd}^{N-1} c_i & c_4 & \cdots & c_1 \\ \sum_{i=0,even}^{N-2} d_i & \sum_{i=1,odd}^{N-1} d_i & d_2 & \cdots & d_{N-1} \\ \sum_{i=0,even}^{N-2} d_i & \sum_{i=1,odd}^{N-1} d_i & d_0 & \cdots & d_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=0,even}^{N-2} d_i & \sum_{i=1,odd}^{N-1} d_i & d_4 & \cdots & d_1 \end{bmatrix}. \tag{4.A.2}$$

We now zero rows 1 through $N/2 - 1$ and rows $N/2 + 1$ through $N - 1$ by adding row $i - 1$ to row i, for $i = N/2 - 1, \ldots, 1$ and $i = N - 1, \ldots, N/2 + 1$. The resulting matrix is given by

$$\mathbf{T}'' = \begin{bmatrix} \sum_{i=0,even}^{N-2} c_i & \sum_{i=1,odd}^{N-1} c_i & c_2 & \cdots & c_{N-1} \\ 0 & 0 & c_0 + c_2 & \cdots & c_{N-3} + c_{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & c_4 + c_6 & \cdots & c_1 + c_3 \\ \sum_{i=0,even}^{N-2} d_i & \sum_{i=1,odd}^{N-1} d_i & d_2 & \cdots & d_{N-1} \\ 0 & 0 & d_0 + d_2 & \cdots & d_{N-3} + d_{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & d_4 + d_6 & \cdots & d_1 + d_3 \end{bmatrix}. \tag{4.A.3}$$

We now apply the same procedure for columns 2 and 3, then for columns 4 and 5, and so on. When the procedures are complete, the resulting matrix needs to be multiplied by a simple permutation matrix to obtain the Hessenberg form. For general N even, the transformation into upper Hessenberg form becomes somewhat more complicated after the first two columns. However, for the special case $N = 2^k$, the transformation matrices $\mathbf{S}$ and $\mathbf{R}$ (Eq. 4.A.1) are given by a closed form formula.

For $N = 2$, the right transformation matrix $\mathbf{R}$ takes the form

$$\mathbf{R}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.A.4}$$

For $N = 2^k$, where $k \geq 2$, $\mathbf{R}_N$ is given by

$$\mathbf{R}_N = \begin{bmatrix} \mathbf{R}_{N/2} & \mathbf{0} \\ \mathbf{R}_{N/2} & \mathbf{R}_{N/2} \end{bmatrix}, \tag{4.A.5}$$

where $\mathbf{0}$ is an $N/2 \times N/2$ matrix of all zeros.

The left transformation matrix $\mathbf{S}$ is defined in terms of Kronecker products. Let

$$\mathbf{S}^n = \mathbf{S} \otimes \mathbf{S} \otimes \ldots \otimes \mathbf{S} \qquad (n \text{ times}) \tag{4.A.6}$$

where $\otimes$ denotes the Kronecker product. For $N = 2$, let

$$\mathbf{S}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{4.A.7}$$

Then for $N = 2^k$ where $k \geq 2$, $\mathbf{S}_N$ is given by

$$\mathbf{S}_N = \begin{bmatrix} \mathbf{S}_2^{log_2 N - 1} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_2^{log_2 N - 1} \end{bmatrix}. \tag{4.A.8}$$

Appendix 4.B

We prove that $\mathbf{B}_N^{-1}$ in Eq. 4.13 is the inverse of the $N \times N$ BFT matrix in Eq. 4.7. Multiplying the two matrices, we get

$$\begin{bmatrix} \mathbf{B}^{ul} & \mathbf{B}^{ur} \\ \mathbf{B}^{urT} & \mathbf{B}^{lr} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{B}^{ul}\mathbf{A} + \mathbf{B}^{ur}\mathbf{C}^T & \mathbf{B}^{ul}\mathbf{C} + \mathbf{B}^{ur}\mathbf{D} \\ \mathbf{B}^{urT}\mathbf{A} + \mathbf{B}^{lr}\mathbf{C}^T & \mathbf{B}^{urT}\mathbf{C} + \mathbf{B}^{lr}\mathbf{D} \end{bmatrix}. \tag{4.B.1}$$

The subscripts that denote the size of the submatrices have been dropped. We note that sizes of the upper left and lower right submatrices are $(N-2) \times (N-2)$ and 2×2, respectively. We now use Eqs. 4.8-4.11 and 4.14-4.16. We have

$$\mathbf{B}^{ul}\mathbf{A} + \mathbf{B}^{ur}\mathbf{C}^T = \begin{bmatrix} 0 & 1 & 0 & \cdots & & & & 0 \\ 0 & 1 & 0 & \cdots & & & & 0 \\ 1 & 1 & 1 & 0 & \cdots & & & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \ddots & & \vdots \\ 0 & 0 & \cdots & & & 0 & 1 & 0 \\ 1 & 1 & 0 & \cdots & & & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \end{bmatrix} = \mathbf{I}_{(N-2)\times(N-2)},$$

where $\mathbf{I}$ is the identity matrix. The upper-right submatrix of Eq. 4.B.1 is

$$\mathbf{B}^{ul}\mathbf{C} + \mathbf{B}^{ur}\mathbf{D} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \overline{\mathbf{B}}_{N-4} & \\ 1 & 1 & & \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = 0 + 0 = 0.$$

The lower-left submatrix can be simplified as

$$\mathbf{B}^{urT}\mathbf{A} + \mathbf{B}^{lr}\mathbf{C}^T = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \end{bmatrix} = 0 + 0 = 0.$$

Finally, the lower-right submatrix takes the form

$$\mathbf{B}^{urT}\mathbf{C} + \mathbf{B}^{lr}\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \mathbf{I}_{2\times 2}.$$

Substituting these results in Eq. 4.B.1 shows that we do indeed have the correct inverse matrix.

Table 4.1: Summary of Coding Results for Letters. Originals: 256x256=8192 Bytes.

Letter	Original Entropy	Original Run-length (Bytes)	Transformed Entropy	Transformed Run-length (Bytes)	Savings
a	0.706	320	0.158	220	45.5%
b	0.753	315	0.168	182	73.1%
c	0.596	257	0.128	186	38.2%
d	0.765	353	0.174	218	61.9%
e	0.630	261	0.137	195	33.9%
m	0.671	262	0.158	181	44.8%
n	0.504	176	0.111	114	54.4%
o	0.660	292	0.147	200	46.0%

Table 4.2: Summary of Coding Results for Fingerprints. Originals: 512x512=32768 Bytes.

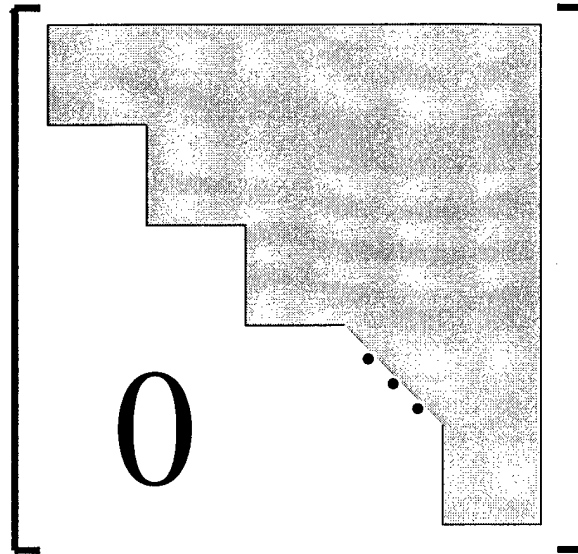| Print	Original Entropy	Original Run-length (Bytes)	Transformed Entropy	Transformed Run-length (Bytes)	Savings
1	0.971	14397	0.537	12615	14.1%
2	0.924	12528	0.477	10977	14.1%
3	0.854	13826	0.499	12104	14.2%
4	0.895	15735	0.522	13754	13.9%
5	0.901	12804	0.478	11202	14.3%
6	0.979	11771	0.513	10227	15.1%

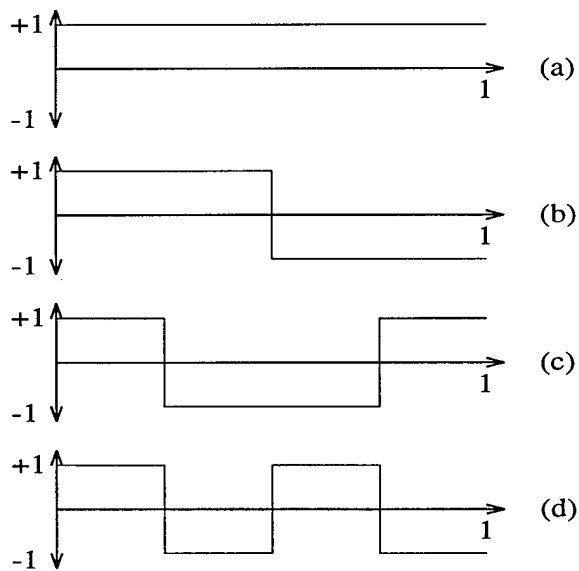Figure 4.1: General form of an upper Hessenberg matrix.

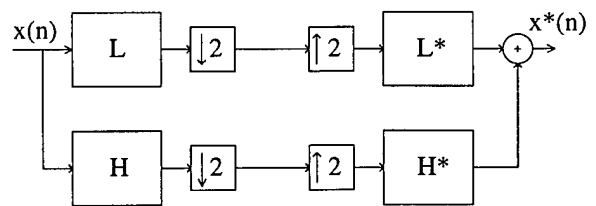Figure 4.2: The first four Walsh waveforms defined over the interval $[0, 1)$.

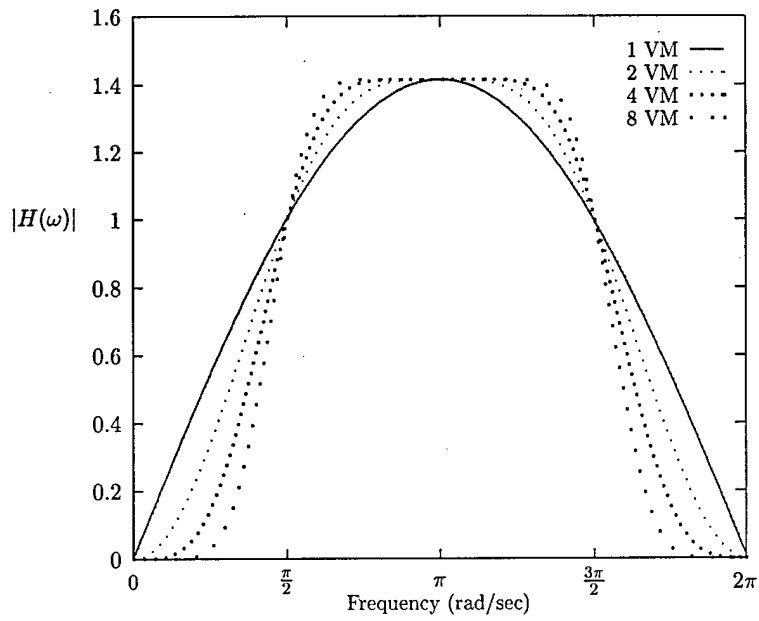Figure 4.3: 2-band perfect reconstruction filter bank.

57

Figure 4.4: Frequency characteristics of Daubechies wavelet filters with 1, 2, 4, and 8 vanishing moments (VM).
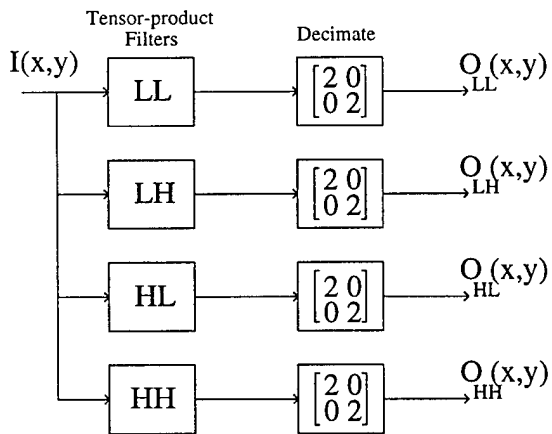

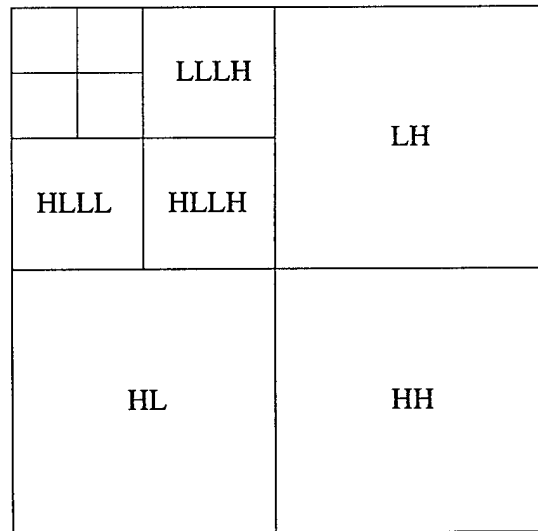Figure 4.5: 2D 2-band perfect reconstruction filter bank.

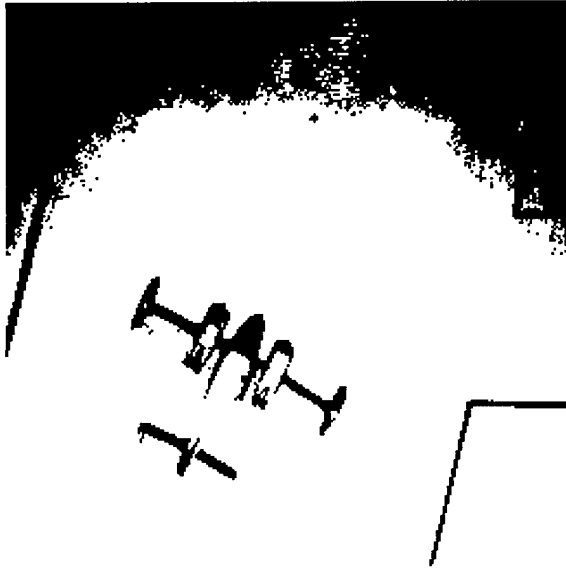Figure 4.6: Sequency regions of 3-stage 2D decomposition.

58

Figure 4.7: Original binary image of plane (256x256).

Figure 4.8: 3-stage decomposition of plane using binary wavelet transform.

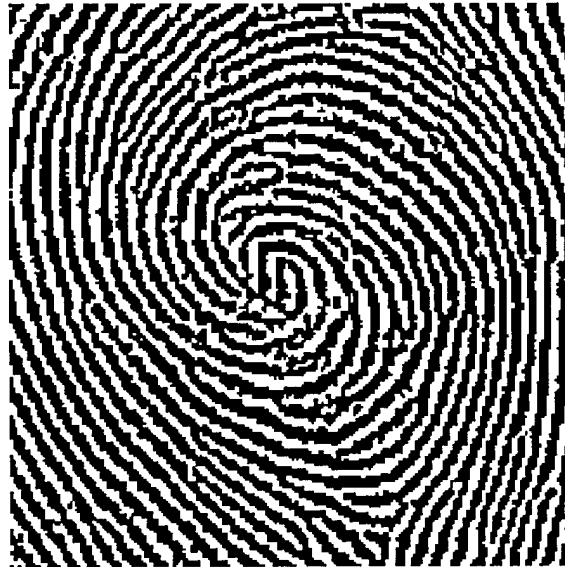Figure 4.9: 3-stage decomposition of plane after thresholding real wavelet transform.

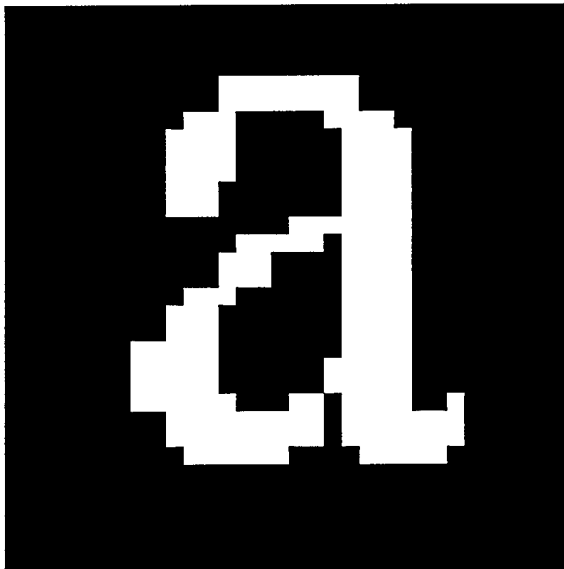Figure 4.10: Example of fingerprint image.

Figure 4.11: Original image of the letter 'a'.

(a)

(b)

(c)

(d)
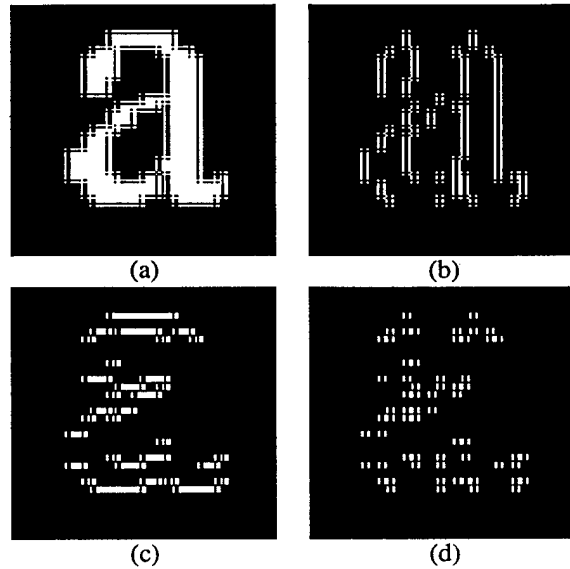
Figure 4.12: Multiresolution images of the letter 'a': (a) LL; (b) LH; (c) HL; (d) HH.
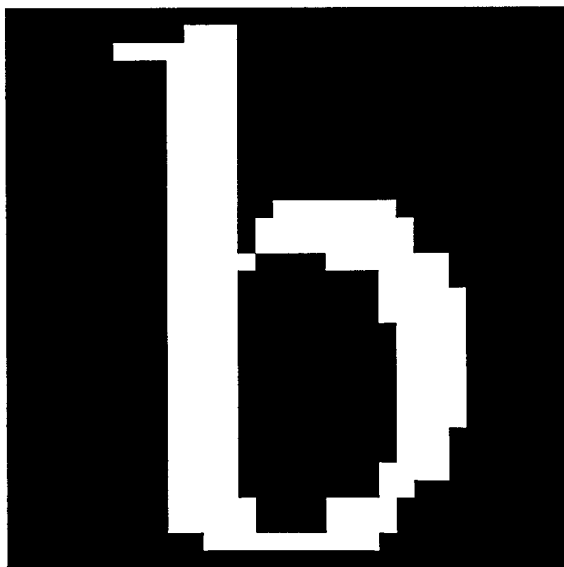

Figure 4.13: Original image of the letter 'b'.
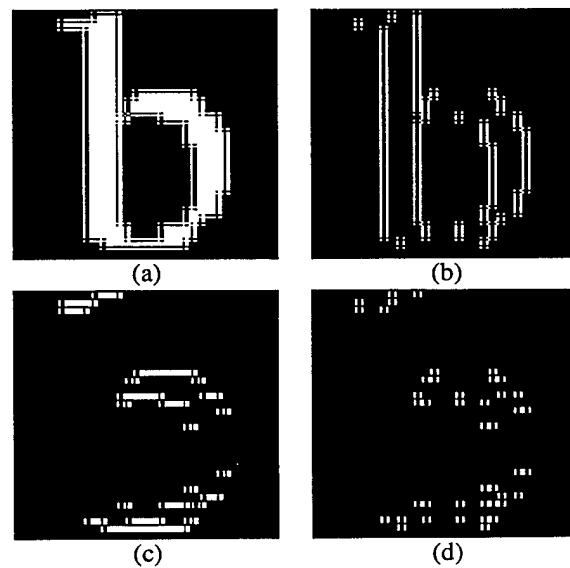

(a)

(b)

(c)

(d)

Figure 4.14: Multiresolution images of the letter 'b': (a) LL; (b) LH; (c) HL; (d) HH.

60

Chapter 5

Personnel and Activities Partially Supported by Grant

5.1 Professional Personnel

- Principal investigators

    1. Prof. A. H. Tewfik
    2. Prof. M. Davidowitz (Year 1 only)

- Graduate students

    1. Srinath Hosur.
    2. Bin Zhu, (Permanent resident).
    3. Mitchell Swanson, (US citizen).
    4. Khalid Hamdy, (US citizen).
    5. Sameh Sowelam.

- Graduate M.Sc. theses completed with partial funding from this grant

    1. Mitchell Swanson, Thesis on "Binary Wavelet Decompositions of Binary Images," Feb. 1995.

- Graduate Ph.D. theses completed with partial funding from this grant

    1. Srinath Hosur, Thesis on "Recursive Matrix Factorization Algorithms in Adaptive Filtering and Mobile Communications", Oct. 1995.
    2. Sameh Sowelam, Thesis on "Optimal Sequential Waveform Selection For Radar Target Imaging and Classification", Feb. 1997.
    3. Mitchell Swanson, Thesis on "Issues in Multimedia Databases", March 1997.

5.2 Publications in technical journals by personnel supported by grant

1. "Fast Multiscale Statistical Signal Processing Algorithms," A. H. Tewfik and M.- Y. Kim, *IEEE Trans. on Signal Proc.* , vol. 42, no. 3, pp. 572-585, March 1994.

2. "Recent Progress in the Application of Wavelets in Surveillance Systems," A. H. Tewfik, S. Hosur and S. Sowelam, invited paper, *Optical Engineering*, vol. 33, no. 8, pp. 2509-2519, August 1994.

3. "Multiscale Difference Equation Signal Models: Part I Theory," M. Ali and A. H. Tewfik, *IEEE Trans. on Signal Proc.* , vol. 43, no. 10, p. 2332-2345, Oct. 1995.

4. "Time Delay Estimation Using Wavelet Transform for PW Ultrasound," X.-L. Xu, A. H. Tewfik and J. F. Greenleaf, *Annals of Biomedical Engineering*, vol. 23, pp. 612-621, 1995.

5. "Completeness and Stability of Partial Wavelet Domain Signal Representations" A. H. Tewfik and H. Zou, *IEEE Trans. on Signal Proc.* , vol. 43, no. 11, pp. 2570- 2581, Nov. 1995.

6. "Efficient design of Ultrasound True-Velocity Flow Mapping," Y. M. Kadah and A. H. Tewfik, *IEEE EMB Mag.*, vol. 15, no. 5, pp. 118-125, Sep-Oct 1996.

7. "Binary Wavelet Decomposition of Binary Images," M. Swanson and A. H. Tewfik, *IEEE Trans. on Image Proc.* , vol. 5, no. 12, pp. 1637-1650, Dec. 1996.

8. "A Wavelet Transform Domain Adaptive Algorithm," S. Hosur and A. H. Tewfik, *IEEE Trans. on Signal Proc.* , March 1997.

9. "Expert Computer Vision Based Crab Recognition System," K. J. Han and A. H. Tewfik, to appear in *IEEE Computer*, 1997.

10. "Arithmetic Coding with Dual Symbol Sets and Its Performance Analysis," submitted to *IEEE Trans. on Image Proc.* , Nov. 1995, revised March 1997.

11. "Multiscale Difference Equation Signal Models: Part II Coding and Decoding Techniques," M. Ali and A. H. Tewfik, submitted to *IEEE Trans. on Signal Proc.* , Nov. 1995, revised March 1997.

12. "ULV and Generalized ULV Subspace Tracking Adaptive Algorithm," S. Hosur, A. H. Tewfik and D. Boley, submitted to *IEEE Trans. on Signal Proc.* , Nov. 1995.

13. "Robust Audio Watermarking Using Perceptual Masking," M. Swanson, B. Zhu, A. H. Tewfik and L. Boney, submitted to special issue of *Signal Processing.*, Jan. 1997.

14. "Waveform Selection in Radar Target Classification," S. Sowelam and A. H. Tewfik, submitted to *IEEE Trans. on Info. Theory*, Feb. 1997.

15. "Waveform Selection for Imaging Unknown Radar Targets," S. Sowelam and A. H. Tewfik, submitted to *IEEE Trans. on Image Processing*, Feb. 1997.

16. "Optimal Waveforms for Wideband Radar Imaging," S. Sowelam and A. H. Tewfik, invited paper, submitted to Special Signal Proc. issue of the *J. of Franklin Institute*, Feb. 1997.

17. "Multiresolution Video Watermarking using Perceptual Models and Scene Segmentation," M. Swanson, B. Zhu, A. H. Tewfik and B. Chau, submitted to special issue of *IEEE J. on Selected Areas in Comm.*, Feb. 1997.

18. "Integrated Image Coding And Content Based Retrieval," M. Swanson and A. H. Tewfik, in preparation, submitted to *IEEE Trans. on Image Proc.* , March 1997.

5.3   Interactions

- Plenary talks, keynote addresses and invited tutorials by personnel supported by grant

1. "Wavelets: Theory and Applications," *1994 IEEE Time-Frequency and Time- Scale Symp.*, Philadelphia, PA, Oct. 1994.

2. "Wavelets: Theory and Applications," *1995 IEEE EMBS Summer School*, Siena, Italy, July 1995.

- Invited papers presented at meetings, conferences and seminars

1. "Recent Developments in Wavelet Theory and Applications," A. H. Tewfik, invited paper, *NSF Sponsored International Conference on Mathematical Analysis and Signal Processing*, Cairo, Egypt, Jan. 1994.

2. "Recent Progress in the Application of Wavelets in Surveillance Systems," A. H. Tewfik and S. Hosur, in *Proc. SPIE Conf. on Wavelet Applications*, SPIE Proc. Vol. 2242, April, 1994.

3. "M-target Class Adaptive Radar Range-Doppler Imaging in Clutter: Theory and Experimental Results," A. H. Tewfik and S. Sowelam, in *Proc. SPIE Conf. on Wavelet Applications for Dual Use*, April 1995.

4. "Wavelets: A Passing Wave or a Truly Useful Tool?," A. H. Tewfik, in *Conf. Digital Processing Technology (Critical Review)*, April 1995.

5. "Coding and Decoding Techniques for Multiscale Difference Equation Signal Models," M. Ali and A. H. Tewfik, *1995 IEEE Workshop on Nonlinear Signal and Image Processing*, Greece, June 1995.

6. "Adaptive Signal Representations in Signal Acquisition and Processing", M. Ali and A. H. Tewfik, *17th Annual Int. Conf. of the IEEE EMBS*, Sept. 1995.

7. "Transparent Robust image watermarking," M. Swanson, B. Zhu and A. H. Tewfik, in Proc. 1996 *IEEE Int. Conf. Image Proc.*, Lausanne, Switzerland, Sept. 1996.

- Papers presented at meetings, conferences and seminars by personnel supported by grant

1. "Perfect Reconstruction Filter Banks with Arbitrary Regularity," A. H. Tewfik, in *Proc. SPIE Conf. on Wavelet Applications*, SPIE Proc. Vol. 2242, April, 1994.

2. "Second Generation Audio Information Coding," A. H. Tewfik, M. Ali and V. Viswanathan, in *Proc. SPIE Conf. on Wavelet Applications*, SPIE Proc. Vol. 2242, April, 1994.

3. "Generalized URV Subspace Tracking LMS Algorithm" S. Hosur, A. H. Tewfik and D. Boley, in *Proc. of the 1994 IEEE Conf. on Acoust. Speech and Signal Proc.*, Adelaide, Australia, April 1994.

4. "Multiscale Difference Equation Signal Modeling and Analysis" M. Ali and A. H. Tewfik, in *Proc. of the 1994 IEEE Conf. on Acoust. Speech and Signal Proc.*, Adelaide, Australia, April 1994.

5. "Wavelet Domain Bearing Estimation in Unknown Correlated Noise" A. H. Tewfik, in *Proc. of the 1994 IEEE Conf. on Acoust. Speech and Signal Proc.*, Adelaide, Australia, April 1994.

6. "ECG Coding by Wavelet Transform Extrema," A. E. Cetin, A. H. Tewfik and Y. Yardimci, *1994 IEEE Symp. Time-Freq. and Time-Scale*, Oct. 1994.

7. "Optimal Waveform Selection in Range-Dopler Imaging," S. Sowelam and A. H. Tewfik, *1994 IEEE Int. Conf. Image Proc.*, Nov. 1994.

8. "Wavelet Decomposition of Binary Finite Images," M. Swanson and A. H. Tewfik, *1994 IEEE Int. Conf. Image Proc.*, Nov. 1994.

9. ""Waveform Selection for High Resolution Range-Dopler Imaging," A. H. Tewfik, *1995 ONR Wideband RF Science and Technology Workshop.*, Jan. 1995.

10. "Low Bit Rate Transparent Image Coding With Optimized Mixed Representations," A. H. Tewfik and B. Zhu, in *Proc. SPIE Conf. on Wavelet Applications for Dual Use*, April 1995.

11. "Image Coding with Mixed Representations and Visual Masking" B. Zhu, A. H. Tewfik and O. Gerek, in *Proc. of the 1995 IEEE Conf. on Acoust. Speech and Signal Proc.*, Detroit, MI, May 1995.

12. "Detection of Weak Signals Using Adaptive Stochastic Resonance" A. Asdi and A. H. Tewfik, *Proc. of the 1995 IEEE Conf. on Acoust. Speech and Signal Proc.*, Detroit, MI, May 1995.

13. "Coding and Decoding Techniques for Multiscale Difference Equation Models," *IEEE Workshop on Nonlinear Signal and Image Processing*, Neos Marmaras, Greece, June 1995.

14. "Space-Invariant True-Velocity Flow Mapping Using Coplanar Observations," Y. M. Kadah and Ahmed H. Tewfik, *17th Annual Int. Conf. of the IEEE EMBS*, Sept. 1995.

15. "Waveform and Beamform Design for Doppler Ultrasound Vector Flow Mapping," Y. M. Kadah and A. H. Tewfik, *17th Annual Int. Conf. of the IEEE EMBS*, Sept. 1995.

16. "Adaptive Multiuser Receiver Schemes for Antenna Arrays," S. Hosur, A. H. Tewfik and V. Ghazi-Moghadam, *Sixth IEEE Int. Symp. On Personal, Indoor and Mobile Radio Comm. (PIMRC'95)*, Toronto, Canada, Sept. 1995.

17. "Theory of True Velocity Duplex Imaging Using A Single Transducer," Yasser M. Kadah and A. H. Tewfik, *1995 IEEE Int. Conf. Image Proc.*, Washington, D.C., Oct. 1995.

18. "Visual Masking and the Design of Magnetic Resonance Image Acquisition," H. H. Garnaoui and A. H. Tewfik, *1995 IEEE Int. Conf. Image Proc.*, Washington, D.C., Oct. 1995.

19. "Image Coding with Wavelet Representations, Edge Information and Visual Masking," B. Zhu, A. H. Tewfik, M. A. Colestock, O. N. Gerek and A. E. Cetin, *1995 IEEE Int. Conf. Image Proc.*, Washington, D.C., Oct. 1995.

20. "True Velocity Estimation Using the Correlation Technique," Y. M. Kadah and Ahmed H. Tewfik, *1995 IEEE Int. Ultrasonics Symp.*, Nov. 1995.

21. "Compact Angular Support Beams for Space Invariant Vector Flow Mapping," Y. M. Kadah and A. H. Tewfik, *1995 IEEE Int. Ultrasonics Symp.*, Nov. 1995.

22. "Efficient Coding of Wavelet Trees and Its Applications in Image Coding," B. Zhu, E. Yang and A. H. Tewfik, *1996 Visual Comm. and Image Proc. (VCIP'96)*, Orlando, Fl, March 1996.

23. "Image Coding for Content Based Retrieval," M. D. Swanson, S. Hosur and A. H. Tewfik, *1996 Visual Comm. and Image Proc. (VCIP'96)* Orlando, Fl, March 1996.

24. "Coding for Content-Based Retrieval", M. Swanson and A. H. Tewfik, in *Proc. of the 1996 IEEE Conf. on Acoust. Speech and Signal Proc.*, Atlanta, GA, May 1996.

25. "High Quality Audio Coding Using Adaptive Signal Representation," K. Hamdy, M. Ali and A. H. Tewfik, in *Proc. of the 1996 IEEE Conf. on Acoust. Speech and Signal Proc.*, Atlanta, GA, May 1996.

26. "Optimal Subset Selection for Adaptive Signal Representation," M. Nafie, M. Ali and A. H. Tewfik, in *Proc. of the 1996 IEEE Conf. on Acoust. Speech and Signal Proc.*, Atlanta, GA, May 1996.

27. "Modeling Techniques for Multiscale Difference Equation Signal Models," M. Ali and A. H. Tewfik, in *Proc. of the 1996 IEEE Conf. on Acoust. Speech and Signal Proc.*, Atlanta, GA, May 1996.

28. "Wavelet transform domain RLS algorithm," S. Hosur and A. H. Tewfik, in *Proc. of the 1996 IEEE Conf. on Acoust. Speech and Signal Proc.*, Atlanta, GA, May 1996.

29. "Digital watermarks for audio signals," L. Boney, A. H. Tewfik and K. Hamdy, in *Proc. IEEE Multimedia Conf.*, Hiroshima, Japan, June 1996.

30. "Digital watermarks for audio signals," L. Boney, A. H. Tewfik and K. Hamdy, in *Proc. of the VII European Signal Proc. Conf. (Eusipco-96)*, Trieste, Italy, Sept. 1996.

31. "Robust Data hiding for images," M. Swanson, B. Zhu and A. H. Tewfik, in *Proc. 1996 IEEE DSP Workshop*, Loen, Norway, Sept. 1996.

32. "Transparent Robust Authentication and Distortion Measurement Technique for Images," B. Zhu, M. Swanson and A. H. Tewfik, in *Proc. 1996 IEEE DSP Workshop*, Loen, Norway, Sept. 1996.

33. "Binary valued wavelet decomposition of binary images," M. Swanson and A. H. Tewfik, in *Proc. of the VII European Signal Proc. Conf. (Eusipco-96)*, Trieste, Italy, Sept. 1996.

34. "Dual Set Arithmetic Coding and its Application in Image Coding," B. Zhu, E. Yang and A. H. Tewfik, in *Proc. of the VII European Signal Proc. Conf. (Eusipco- 96)*, Trieste, Italy, Sept. 1996.

35. "Adaptive waveform selection for target classification," S. Sowelam and A. H. Tewfik, in *Proc. of the VII European Signal Proc. Conf. (Eusipco-96)*, Trieste, Italy, Sept. 1996.

36. "Embedded object dictionaries for image database browsing and searching," M. Swanson and A. H. Tewfik, *1996 IEEE Int. Conf. Image Proc.*, Lausanne, Switzerland, Sept. 1996.

37. "Expert computer vision based crab recognition system," K. Han and A. H. Tewfik, *1996 IEEE Int. Conf. Image Proc.*, Lausanne, Switzerland, Sept. 1996.

- Consultative and advisory functions to DoD
  - Visit to MIT Lincoln Laboratories, Nov. 1993.
  - Visit to NAWCWPNS, China Lake, CA, Dec. 1993.
  - Visit to NRaD, SanDiego, CA, Jan. 1995.

- Patents
  Two filed in 1996.

Bibliography

[1] H. L. Van Trees, *Detection, Estimation and Modulation Theory*, New York, Wiley, 1968, Part III.

[2] H. Naparst, "Dense Target Signal Processing," *IEEE Trans. Info. Theory*, vol. IT-37, no. 2, pp. 317-327, 1991.

[3] P. Maass, "Wideband Approximation and Wavelet Transform," in *Radar and Sonar*, A. Grünbaum, M. Bernfeld, and R. Blahut, eds., IMA Volume in Mathematics and its Applications, Springer-Verlag, New York, 1991.

[4] C.H. Wilcox, "The Synthesis Problem for Radar Ambiguity Functions," in *Radar and Sonar*, A. Grünbaum, M. Bernfeld, and R. Blahut, eds., IMA Volume in Mathematics and its Applications, Springer-Verlag, New York, 1991.

[5] N.F. Ezquerra, "Target Recognition Considerations," in *Principles of Modern Radar*, J. L. Eaves and E.K. Reedy, eds., New York: Van Nostrand Reinhold, 1987.

[6] D. Guili, "Polarization Diversity in Radar," *Proc. IEEE*, vol. 74 pp.245-269, Feb. 1986.

[7] S. R. Cloude, "Polarimetric Techniques in Radar Signal Processing," *Microwave Journal*, vol. 26, no. 7, July 1983, pp. 119-127.

[8] W. A. Holm, "Polarimetric Fundamentals and techniques," in *Principles of Modern Radar*, J. L. Eaves and E.K. Reedy, eds., New York: Van Nostrand Reinhold, 1987.

[9] D. L. Moffat *et al.*, "Transient Response Characteristics in Identification and Imaging," *IEEE Trans. Antennas Propagat.*, vol. 29, no. 29, March 1981, pp. 192-205. pp. 893-901, August 1965.

[10] D. L. Moffat *et al'*, "Transient Response Characteristics in Identification and Imaging," *IEEE Trans. Antennas Propagat.*, vol. AP-29, pp. 192-205, March 1981.

[11] C.E. Baum *et al.*, "The Singularity Expansion Method and Its Application to Target Identification," *Proceedings of the IEEE*, vol. 79, no. 10, pp. 1481-1491, October 1991.

[12] L. C. Potter and R. L. Moses, "Attributed Scattering Centers for SAR ATR," *IEEE Trans. Image Processing*, vol. 6, no. 1, January 1997, pp. 79-91.

[13] T. L. Grettenberg, "Signal Selection in Communication and Radar Systems," *IEEE Trans. Inform. Theory*, vol. IT-9, pp. 265-272, Oct. 1963.

[14] S. Kullback and R. A. Leibler, "On Information and Sufficiency," *Ann. Math. Stat.*, vol.22, pp.79-86, 1951.

[15] R. E. Blahut, *Principles and Practice of Information Theory*, Addison-Wesley, 1987.

[16] H. Chernoff, *Sequential Analysis and Optimal Design*, Society for Industrial and Applied Mathematics, Philadelphia, 1972.

[17] E. J. Kelly and R.P. Wishner, "Matched-Filter Theory for High-Velocity Accelerating Targets," *IEEE Trans. Military Electronics*, vol. MIL-9, Jan. 1965, pp. 56-69.

[18] S. Sowelam and A.H. Tewfik, "Optimal Waveform Selection in Range-Doppler Imaging," *Proc. of the Intl. Conf. on Image Processing*, Austin, TX, pp. 441-445.

[19] A.H. Tewfik, S. Hosur and S. Sowelam, "Recent Progress in the Application of Wavelets in Surveillance Systems," *Optical Engineering*, vol. 33, no. 8, pp. 2509-2519, 1994.

[20] M. Soumekh, "Band-Limited Interpolation from Unevenly Spaced Sampled Data," *IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-36, p. 110-122*, 1988.

[21] C.W. Therrien, "A Sequential Approach to Target Discrimination," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-14, no.3, pp. 433-440, May 1978.

[22] A. Wald, *Sequential Analysis*, John Wiley & Sons, New York, 1947.

[23] J. J. Bussgang and D. Middleton, "Optimum Sequential Detection of Signals in Noise," *IRE Trans. on Inform. Theory*, vol. IT-1, pp. 5-18, 1955.

[24] S. Sowelam and A.H. Tewfik, "Adaptive Waveform Selection for Target Classification", *Proc. of EUSIPCO-96, VIII European Signal Proc. Conf.*, Trieste, Italy, September, 1996.

[25] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley, New York, 1993.

[26] G. E. Shilov, *Linear Algebra*, Prentice Hall, 1971.

[27] A. C. Atkinson and V. V. Federov, "Optimal Design: Experiments for Discriminating Between Several Models", *Biometrika*, vol. 62, pp. 289-303, 1975.

[28] A. E. Cawkell, "Picture queries and picture databases," *Journal of Info. Sciences*, vol. 19, pp. 409–423, 1993.

[29] R. Jain, A. Pentland, and D. Petkovic, "Workshop Report: NSF-ARPA Workshop on Visual Information Management Systems." Available on the WWW at http://vision.ucsd.edu/papers/vimsReport95.html, June 1995.

[30] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, "The QBIC Project: Querying Images by Content using Color, Texture, and Shape," *SPIE 1993 Storage and Retrieval for Image and Video Databases*, vol. 1908, pp. 173–187, 1993.

[31] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC System," *Computer*, vol. 28, pp. 23–32, Sept. 1995.

[32] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Content-Based Manipulation of Image Databases." MIT Media Lab Tech Report No. 255, 1993.

[33] G. Wallace, "The JPEG Still Picture Compression Standard," *Communications of the ACM*, vol. 34, no. 4, pp. 30–44, 1991.

[34] C. E. Jacobs, A. Finkelstein, and D. H. Salesin, "Fast Multiresolution Image Querying," in *Proc. ACM SIGGRAPH Conference on Computer Graphics*, (Los Angeles, CA), pp. 277–286, 1995.

[35] B. S. Manjunath and W. Y. Ma, "Browsing Large Satellite and Aerial Photographs," in *Proc. IEEE Int. Conf. on Image Proc.*, (Lausanne, Switzerland), pp. 765–768, 1996.

[36] K. Hirata and T. Kato, "Rough Sketch-Based Image Information Retrieval," *NEC Research & Development*, vol. 34, no. 2, pp. 263–273, 1993.

[37] J. Gary and R. Mehrotra, "Shape Similarity-Based Retrieval in Image Database Systems," *SPIE Image Storage and Retrieval Systems*, vol. 1662, pp. 2–8, 1992.

[38] H. C. Lin, L. L. Wang, and S. N. Yang, "Color Image Retrieval Based on Hidden Markov Models," in *Proc. 1995 IEEE Int. Conf. on Image Proc.*, (Washington, DC), pp. 342–345, 1995.

[39] S. F. Chang, "Compressed-Domain Techniques for Image/Video Indexing and Manipulation," in *Proc. 1995 IEEE Int. Conf. on Image Proc.*, (Washington, DC), pp. 314–317, 1995.

[40] J. Li, S. Kim, and C.-C. J. Kuo, "Focus of Attention (FOA) Identification from Compressed Video for Automatic Target Recognition (ATR)," in *Proc. 1995 IEEE Int. Conf. on Image Proc.*, (Washington, DC), pp. 342–345, 1995.

[41] M. D. Swanson, S. Hosur, and A. H. Tewfik, "Coding for Content-Based Retrieval," in *Proc. 1996 IEEE Int. Conf. on Acoustics, Speech and Signal Proc.*, vol. IV, (Atlanta, GA), pp. 1959–1962, 1996.

[42] M. D. Swanson and A. H. Tewfik, "Embedded Object Dictionaries for Image Database Browsing and Searching." in *Proc. 1996 Int. Conf. on Image Proc.*, Sept. 1996.

[43] M. Vetterli, *Wavelets and Subband Coding.* Englewood Cliffs, NJ: P T R Prentice-Hall, Inc., 1995.

[44] A. Gersho and R. Gray, *Vector Quantization and Signal Compression.* Boston, MA: Kluwer Academic, 1992.

[45] D. A. Knuth, *The Art of Computer Science: Sorting and Searching*, vol. 3. Reading, MA: Addison Wesley, 1973.

[46] B. Zhu, A. Tewfik, and O. Gerek, "Low Bit Rate Near-Transparent Image Coding," in *Proc. of the SPIE Int. Conf. on Wavelet Apps. for Dual Use*, vol. 2491, (Orlando, FL), pp. 173–184, 1995.

[47] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Trans. on Patt. Anal. and Machine Intell.*, vol. 14, pp. 710–732, July 1992.

[48] T. Parks and R. G. Shenoi, "Time-Frequency Basis Functions," in *Proc. 1990 IEEE Int. Conf. on Acoust. Speech and Signal Proc.*, (Albuquerque, N.M.), pp. 2459–2462, 1990.

[49] A. Lewis and G. Knowles, "Image Compression Using the 2-D Wavelet Transform," *IEEE Trans. on Image Processing*, vol. 1, pp. 244–250, Apr. 1992.

[50] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients," *IEEE Trans. on Signal Processing*, vol. 41, pp. 3445–3462, Dec. 1993.

[51] Y. Xu, J. B. Weaver, D. M. Healy, Jr., and J. Lu, "Wavelet Transform Domain Filters: A Spatially Selective Noise Filtration Technique," *IEEE Trans. on Image Processing*, vol. 3, pp. 747–758, Nov. 1994.

[52] D. L. Donoho and I. M. Johnstone, "Ideal Spatial Adaptation via Wavelet Shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[53] G. W. Wornell and A. V. Oppenheim, "Wavelet-Based Representations for a Class of Self-Similar Signals with Application to Fractal Modulation," *IEEE Trans. on Info. Theory*, vol. 38, pp. 785–800, Mar. 1992.

[54] G. C. Freeland and T. S. Durrani, "IFS Fractals and the Wavelet Transform," in *Proc. 1990 IEEE Int. Conf. on Acoust. Speech and Signal Proc.*, (Albuquerque, NM), pp. 2345–2348, 1990.

[55] P. Moulin, "Wavelet Thresholding Techniques for Power Spectrum Estimation," *IEEE Trans. on Signal Processing*, vol. 42, pp. 3126–3136, Nov. 1994.

[56] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," *IEEE Signal Processing Mag.*, vol. 8, pp. 14–38, Oct. 1991.

[57] C. Herley, J. Kovacevic, K. Ramchandran, and M. Vetterli, "Tilings of the Time-Frequency Plane: Construction of Arbitrary Orthogonal Bases and Fast Tiling Algorithms," *IEEE Trans. on Signal Processing*, vol. 41, pp. 3341–3359, Dec. 1993.

[58] A. H. Tewfik, S. Hosur, and S. Sowelam, "Recent Progress in the Applications of Wavelets in Surveillance Systems," *Optical Engineering*, vol. 33, pp. 2509–2519, Aug. 1994.

[59] P. Moulin, J. A. O'Sullivan, and D. L. Snyder, "A Method of Sieves for Multiresolution Spectrum Estimation and Radar Imaging," *IEEE Trans. on Info. Theory*, vol. 38, pp. 801–813, Mar. 1992.

[60] P. P. Vaidyanathan, "Unitary and Paraunitary Systems in Finite Fields," in *Proc. 1990 IEEE Int. Symp. on Circuits and Systems*, (New Orleans, LA), pp. 1189–1192, 1990.

[61] S. M. Phoong and P. P. Vaidyanathan, "Paraunitary Filter Banks over Finite Fields," *Preprint*, Oct. 1995.

[62] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, N.J.: P T R Prentice-Hall, Inc., 1992.

[63] G. Caire, R. L. Grossman, and H. V. Poor, "Wavelet Transforms Associated with Finite Cyclic Groups," *IEEE Trans. on Info. Theory*, vol. 39, pp. 1157–1166, July 1993.

[64] K. Flornes, A. Grossmann, M. Holschneider, and B. Torresani, "Wavelets on Finite Fields," *Preprint*, Nov. 1993.

[65] R. E. Blahut, *Algebraic Methods for Signal Processing and Communications Coding*. New York: Springer-Verlag, Inc., 1992.

[66] J. Hong and M. Vetterli, "Hartley Transforms Over Finite Fields," *IEEE Trans. on Info. Theory*, vol. 39, pp. 1628–1638, Sept. 1993.

[67] H. F. Harmuth, *Transmission of Information by Orthogonal Functions*. New York: Springer-Verlag, Inc., 1972.

[68] N. Ahmed and K. R. Rao, *Orthogonal Transforms for Digital Signal Processing*. New York: Springer-Verlag, Inc., 1975.

[69] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Comm. Pure and Applied Math.*, vol. 41, pp. 909–996, Nov. 1988.

[70] R. R. Coifman and M. V. Wickerhauser, "Entropy-Based Algorithms for Best Basis Selection," *IEEE Trans. on Info. Theory*, vol. 38, pp. 713–718, Mar. 1992.

[71] P. Wunsch and A. F. Laine, "Wavelet Descriptors for Multiresolution Recognition of Handprinted Characters," *Pattern Recognition*, vol. 28, no. 8, pp. 1237–1249, 1995.